DOCUMENT RESUME

ED 143 499 SE 022 850

AUTHOR	Lucas, A. M., Ed.; Power, Colin, N., Ed.
TITLE	Research in Science Education, Volume 5. Proceedings of the Annual Conference of the Australian Science Education Research Association (6th, Flinders University, Bedford Park, South Australia, May 19-21, 1975).
INSTITUTION	Australian Science Education Research Association.
PUB DATE	75
NOTE	206p.; For Volume 6, see SE 022 891; Contains occasional small print
EDRS PRICE	MF-$0.83 HC-$11.37 Plus Postage.
DESCRIPTORS	*College Science; *Conference Reports; *Educational Research; *Elementary School Science; Elementary Secondary Education; Higher Education; Instruction; *Science Education; Science Teachers; Secondary School Science
IDENTIFIERS	Australian Science Education Research Association; *Research Reports

ABSTRACT

This volume contains papers presented at the sixth Annual Conference of the Australian Science Education Research Association (ASERA) held at Flinders University in May, 1975. Paper topics include: pupil learning and classroom climate, teacher structuring behavior, the Australian Science Education Project (ASEP), cognitive preference and achievement, student expectations, medical education, biometrics in environmental science, the Keller plan, scientific methodology, perceived need satisfaction, use of learning hierarchies, grade placement of science concepts in primary school science, and the primary science program in New Guinea. (SL)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED143499

RESEARCH IN SCIENCE EDUCATION

Volume 5

Colin N. Power

Australian Science Education Research Association

RESEARCH

IN

SCIENCE EDUCATION

VOLUME 5

Proceedings of the Sixth Annual Conference of the Australian Science Education Research Association, Flinders University, Bedford Park, S.A., May 19–21, 1975.

Edited by: *A. M. Lucas*

General Editor: *Colin N. Power*

AUSTRALIAN SCIENCE EDUCATION RESEARCH ASSOCIATION

All correspondence concerning this publication should be addressed to Colin N. Power, Department of Education, University of Queensland, St. Lucia, 4067, Australia.

Copyright © 1975 by Australian Science Education Research Association.
Published by Australian Science Education Research Association.
Printed by the University of Queensland Printery.

CONTENTS

	Page
Preface	(v)
Barry J. Fraser	
The impact of ASEP on pupil learning and classroom climate	1
Colin N. Power and R. P. Tisher	
Experimental studies of teacher structuring behaviour in ASEP classrooms	13
Jeff R. Northfield	
Changing perspectives of science teacher education that may be associated with curriculum development	23
John M. Owen	
A review of current and planned research concerning the Australian Science Education Project	33
Peter Fensham and Kaye Nickless	
Detectives are born, not made *or* Two sides of a chemistry learning coin	43
Lindsay D. Mackay	
Cognitive preferences and achievement in physics, chemistry, science and mathematics	49
Janet V. Keightley and Effice D. Best	
Student expectations of grade 11 biology in some South Australian schools	59
W.C. Hall and J. Lawrence	
An evaluation of the introductory clinical skills component of the 3rd year medical course at Adelaide University	69
P.R.B. Blood and A.H.A. Bensink	
An example of the role of biometrics in teaching practical environmental science	77
B.G. Collins	
Biology and the Keller plan	87

	Page
B.C. Newman	
Scientific methodology: A Feyerabendian view	99
Lindsay D. Mackay and Jeff R. Northfield	
Validity of perceived need satisfaction as a curriculum evaluation criterion	113
R.L. Stanton and M.N. Maddock	
Science Education and the Supernatural in Papua New Guinea	123
R. Trembath and R.T. White	
Use of learning hierarchies in promoting mastery learning	135
Brian W. Carss	
Thematic structural analysis applied to conceptual hierarchies	143
E. Jungwirth	
Caveat Mentor – Let the teacher beware	153
P.L. Gardner	
Logical connectives in science: A preliminary report	161
John W. Fisher and Robert L. Kimber	
Grade placement of physical and chemical change concepts in primary school science	177
M.N. Maddock	
Has Papua New Guinea's three phase primary science program had any impact at secondary level?	189
J.T. McArthur, B.J. Fraser and L.H.T. West	
Some issues in the simulation of ANOVA by multiple regression techniques	199
Contributors	209

PREFACE

This volume contains the papers presented at the sixth Annual Conference of the Australian Science Education Research Association held at Flinders University in May 1975. The papers reflect the diversity of interests of ASERA members; ranging from philosophy of science and science teaching to experimental studies in learning theory; from primary school science to preparation of doctors and agricultural scientists; and from studies based on local curriculum developments to investigations conducted overseas. The authors do not, therefore, share common research traditions; the commonality is provided by the subject area of the curriculum being investigated within these diverse paradigms.

The standard of the papers is encouraging. ASERA exists to provide a forum for describing work in progress; for discussing methodological issues of importance to the members; and for reporting work of interest mainly to Australian science educators and thus, inappropriate for many foreign journals. These objects are met, in part, by accepting papers for the conference without applying any "quality" criteria. Thus, the papers in the proceedings have not been subject to judgement by an editorial panel, although some were revised in the light of discussion after their presentation; yet few are below the standards of appropriate refereed journals.

The stimulus to research provided by the introduction of the Australian Science Education Project is evident again in this issue; but the interest of members in tertiary teaching/learning is also apparent. The latter is possibly an artifact stemming from the active soliciting of papers in this area by the conference, who also took other decisions that would invalidate studies of trends in Australian science education research based on comparison of this issue with early proceedings: a limit was placed on the number of papers associated with any single contributor, review articles were not accepted, and the short papers used as stimuli for workshop discussion groups have not been included here. The omission of the workshop papers was a deliberate decision of the conference convenors and is not a reflection on their importance; members' were told that the papers and discussion were not for publication in the belief that awareness of ultimate publication may have encouraged unproductive caution.

Professor Tisher has relinquished the position of general editor of the ASERA series. He has had a heavy responsibility, editing the first three volumes alone, and acting as general editor for Volume 4, being responsible for seeing the work through the press. This task was more difficult for him in his new position, at Monash, than it had been previously when he was at the University of Queensland. The Association greatly appreciates the extensive ground work he laid as general editor; as members we look forward to his continued participation in conferences and thank him for his important role in establishing the Association and its publication.

I thank Effie Best and Ian Walker for their assistance in planning the conference reported here.

A.M. Lucas
Flinders University
11 August, 1975

THE IMPACT OF ASEP ON PUPIL LEARNING AND CLASSROOM CLIMATE

Barry J. Fraser

In this paper, two matters will be considered. Firstly, attention will be given to a theoretical model useful for the conceptualization of an evaluation of the Australian Science Education Project (ASEP). Secondly, a number of the experimental findings related to pupil learning and classroom climate will be reported.

Walberg's model for research on instruction

Walberg (1970 & 1971) has proposed the following quasi-functional relationship between learning and three other construct domains, instruction, aptitude and environment:

$$L_h = f(I_i, A_j, E_k)$$

In this model, L_h stands for a multidimensional set of learning outcomes such as skills, understandings and attitudes. I_i represents instruction and could include curricular materials, teaching methods, or media. A_j represents individual pupil aptitudes such as socioeconomic status (SES), intelligence, sex or personality. E_k represents the learning environment and might include class size, classroom climate and certain teacher characteristics.

The equation above may be expanded in the following way to yield seven sets of predictors of learning

$$L_h = f_1(I_i) + f_2(A_j) + f_3(E_k) + f_4(I_i A_j) + f_5(I_i E_k) + f_6(A_j E_k) + f_7(I_i A_j E_k)$$

In this equation, four terms are fairly well represented in the educational research literature; f_1 includes classical experimental studies comparing alternative treatments; f_2 includes correlational studies of prediction and selection; f_3 includes stimulation and enrichment studies; f_4 includes aptitude-treatment interaction studies. On the other hand, Walberg (1970, p. 188) has observed that the terms f_5, f_6 and f_7 have received scant attention from the researcher. Furthermore, it should be noted that any one term in the model may contain several operational representations of both main effects and interactions. For example, A_j could represent a set containing age, sex and the interaction of age with sex.

The terms $f_1(I_i)$ and $f_2(A_j)$ in Walberg's model correspond to two traditions in research identified by Cronbach (1957). f_1 would correspond to the *experimental tradition*, characterized by tight control and the manipulation of variables; for the experimentalist, individual differences would be a nuisance. On the other hand, f_2 would correspond to the *correlational tradition* in which research is naturalistic and nonmanipulative; it is treatment variation that would be an annoyance in correlational research. Cronbach (1957, p. 671) has described these two traditions as "two historic streams of method, thought, and affiliation which run through the last century of our science" and which are quite distinct in "philosophical underpinnings, methods of enquiry, topical interests and loci of application". Further discussion of these two traditions may be found in Cronbach (1957), Walberg (1971) or Cattell (1966).

In the present study, some confluence of experimental and correlational traditions is

essential because of the inclusion of both instructional and aptitudinal variables in the research. Furthermore, in research consistent with Walberg's model, the inclusion of environmental variables introduces further complications. Later discussion will therefore focus on some of the conflicts and compromises involved in incorporating I_t, A_t and E_k variables within the same study.

ASEP evaulation and Walberg's model – Walberg's model highlights three important features relevant in an evaluation of ASEP. Firstly, since environment is a potent influence on learning, it would be desirable that ASEP be evaluated in a naturalistic school setting. Secondly, the model captures something of the complexity of factors operating in school learning studies. Consequently, naive questions about whether ASEP is superior to alternatives would be replaced by more complex questions about which curriculum is better for which learning outcomes for what sorts of pupils in what sorts of learning environments. Thirdly, since ASEP materials were intended to be suitable for all pupils in all schools (ASEP, 1974), the interaction of curriculum with pupil aptitude, I_tA_t, and the interaction of curriculum with learning environment, I_tE_k, must be considered important aspects in evaluating ASEP.

Design of the study

An important step in the present study was the choice of variables and measures for the I_t, I_t, A_t and E_k construct domains of Walberg's model. The vital task of selecting and validating measures of 17 learning outcomes, I_t, has been discussed in detail in Fraser (1974). This battery of 17 instruments, listed later in Table 1, consisted of measures of nine enquiry skills, three aspects of understanding of the nature of science and five attitudinal aims. For this battery of tests the Cronbach α reliabilities ranged from 0.55 to 0.85 with a median of 0.69. This battery of tests was administered as pre-tests in March and again as post-tests in October.

For instruction, I_t, two loosely defined treatment groups were employed. An experimental group used ASEP materials almost exclusively during a school year, with the choice, sequence and pacing of units being decided by the school. The control group studied any non-ASEP science program that schools wished to offer.

For reasons discussed by Fraser and Fisher (1974), the A_t aptitude variables chosen were SES, IQ and sex and the four interactions possible between the three variables. While SES was measured using the scale of Congalton (1969), IQ was measured using a modified Otis verbal test. Furthermore, while a number of E_k variables were included in the study, space stringency has precluded their inclusion in the present paper.

The sample consisted of 1,158 Form 1 pupils from 46 classes, each in a different Melbourne coeducational high school. This number of schools was chosen on the basis of power calculations performed using data from a pilot study in Tasmania (Fraser & Fisher, 1974). Of the 46 classes, 20 comprised the ASEP group while the other 26 formed the control group. It should be pointed out that randomisation was used neither in choosing classes nor in allocating them to alternative treatments. Instead, schools were invited to participate in the study by offering either an ASEP or a control class depending on the school's own preference. Some of the ramifications of the absence of randomisation will be outlined later.

Choice of sampling units. For research conceptualized in terms of Walberg's model, the choice of an appropriate sampling unit highlights some of the conflicting demands, described earlier, in incorporating instructional, aptitudinal and environmental variables within an experiment. Below, four alternative sampling strategies are discussed in terms of their potential usefulness in research

conducted within Walberg's framework.

The first sampling procedure would be to choose individual pupils out of their ordinary classes and then to re-group them for the purposes of the study. Two objections immediately arise. Firstly, such a method is grossly impractical in terms of the disruption caused to teachers and schools. Secondly, sampling individuals destroys several aspects of the natural environment in which science, including ASEP, is normally taught.

A second sampling strategy, which is well documented in Campbell and Stanley (1963) and Peckham, Glass and Hopkins (1969), is to select intact classes and then employ intact class means as the units of statistical analysis. This sampling plan not only overcomes the previous two problems of administrative inconvenience and distortion of the natural learning environment, but also allows the use of the appropriate error term in the statistical analysis (Lindquist, 1953, p. 173). However, while the use of intact class means is ideal for comparing treatments within the experimental tradition, it is quite unsatisfactory for correlational research related to individual differences. Clearly, as Gardner (1975b) and Peckham *et al.* (1969) have pointed out, intact class means cannot reflect individual differences in aptitudes.

In an attempt to overcome some of the problems raised above, researchers often seek a compromise by using a third sampling plan of initially selecting intact classes then employing the individual as the unit for statistical analysis. This method has the advantage of allowing the investigation of individual pupil aptitudes, as well as being convenient and maintaining the ecology of the classroom. One obvious drawback, however, is that this strategy would yield an inappropriate error term and an accompanying inflated F value. But a second and more vital problem, related to confounding between instructional treatment and learning environment, also exists. In order to clarify the nature of this confounding, a brief example will be introduced.

Consider a study in which the relationship between learning and the pupil aptitude, SES, was being investigated using a sample of two classes of 30 pupils, each from a different school. Furthermore, suppose that, when a median split of SES was taken for the entire sample, the distribution of pupils of high and low SES between the two schools was that shown in Figure 1a. When an attempt is made to compare the learning of high SES pupils with that of low SES pupils, serious confounding arises because the high SES group contains a predominance of pupils from School A while the low SES group contains a predominance of pupils from School B. But, Coleman (1966) has pointed out that, in general, schools holding pupils of higher aptitudes tend to provide better learning conditions. In the present example, therefore, it would be impossible to know whether any observed differences in learning should be attributed to SES or to differences between the schools. When this example is extended to larger numbers of aptitudinal variables and schools, similar confounding would arise since pupils of different aptitudes would not be uniformly distributed amongst different schools.

(a) Individual as sampling unit

(b) Class aptitude cell as sampling unit

FIGURE 1: Number of sampling units in each category

In terms of Walberg's model, therefore, sampling intact classes then using individuals as the units of analysis must be deemed unsatisfactory because it gives rise to confounding between pupil aptitude and learning environment. An ideal sampling unit, therefore, would need to capture variation due to both individual pupil aptitude and classroom environment.

In the present study, confounding between instruction and environment was minimized by the use of a fourth sampling strategy which involved selecting intact classes and then forming sampling units by dividing each class into subgroups homogeneous with respect to aptitude. In forming these subgroups, which will be referred to as *class aptitude cells*, use was made of the three aptitudes, SES, IQ and sex, which had been chosen for inclusion in the study. By dichotomizing pupils according to SES, IQ and sex – using a median split of SES and IQ across the whole sample – eight class aptitude cells were generated for each class in the sample.

The way that the use of class aptitude cell sampling units can remove confounding between I_i and E_k is shown in Figure 1b. In the example considered previously, School A and School B would each provide eight class aptitude cells, of which four would be classifiable as high SES and four as low SES. That is, within this sampling plan, the original predominance of School A in the high SES group and of School B in the low SES group are removed. Instead, each school would provide an equal number of high and low SES sampling units.

Furthermore, as well as providing substantial control over confounding between aptitude and environment, the use of class aptitude cell sampling units has several other advantages. Firstly, the initial sampling of intact classes is convenient and preserves the natural environment. Secondly, the existence of eight aptitudinal levels within each class allows the investigation of the relationship of aptitude with learning. Thirdly, the use of class aptitude cells would involve a more appropriate error term, and therefore a less inflated F value, than the alternative strategy of employing individuals as units of analysis.

Confounding between terms in Walberg's model. An incisive conceptual analysis of potential confounding between variables operating in the present experiment may be made within the framework of Walberg's model. Figure 2 illustrates three possible sources of confounding, between aptitude A_j and environment E_k, between instruction I_i and aptitude A_j, and between I_i and E_k.

(a) Between aptitude and environment

(b) Between instruction and aptitude

(c) Between instruction and environment

FIGURE 2: Potential confounding between terms in Walberg's model

The confounding that arises in Figure 2a, between aptitude and environment, has already been discussed in relation to the choice of sampling units. In the present study, substantial control over confounding between A_j and E_k would be achieved by the use of class aptitude cell sampling units.

The second potential source of confounding, between instruction I_t and aptitude A_f, is shown diagramatically in Figure 2b. Since classes were not randomly allocated to treatments, it would be difficult to know whether any observed differences should be attributed to the treatments *per se*, or rather, to possible differences existing between the aptitudes of ASEP and control pupils. This possibility was investigated empirically by comparing the ASEP group with the control group on all 17 pre-test measures and on the three aptitudinal variables, SES, IQ and sex. The fact that no significant differences between ASEP and control groups were found on any of these 20 variables provides some evidence that sampling was "random in effect" (Lord, 1963, p. 38), with respect to pupil aptitudes, and that, therefore, there would be little serious confounding between I_t and A_f in the present study.

The possibility of the confounding between instruction I_t and environment E_k shown in Figure 2c, however, cannot be dismissed, especially since classes were not randomly allocated to instructional treatments. Any observed differences between ASEP and control, therefore, might reflect not only differences in treatments, but also the fact that teachers and schools choosing to study ASEP may have been different in some important respects from those electing to follow traditional programs. Although ASEP and control teachers were not found to be significantly different with respect to sex and years of teaching experience, the mere fact that some teachers *chose* to use ASEP while others *chose* to follow alternative programs suggested that confounding between I_t and E_k was a possibility and that some caution should be exercised in interpreting results.

Of course, the degree of confounding between I_t and E_k might have been reduced by random allocation of classes to treatments. But, while this procedure would enhance the *internal validity*, two problems arise. Firstly, because of financial considerations, it was not feasible to ask schools to use particular curricular materials. Secondly, since it is part of the naturalistic environment that schools choose what program they wish to follow, removal of this choice would adversely affect the *ecological validity* discussed by Bracht and Glass (1968). In studies like the present one, therefore, "teacher and method are confounded in nature" (Snow, 1974, p. 282). For this reason, Snow advocates that it is "preferable to match teachers with congenial treatments" but cautions that "generalizations will be limited by the confounding of teacher and method". Further discussion of aspects of representativeness of educative conditions may be found in Pereboom (1971), Brunswik (1956) and Barker (1968).

Results

The impact of ASEP on learning outcomes. Table 1 contains F values from multiple regression analyses for the contribution made by curricular treatment, pupil aptitude and aptitude-treatment interactions to the prediction of 17 learning outcomes. The second column in Table 1 contains the results for I_t as predictor; that is, for differences between the ASEP and control group. The third column shows F values for A_f, a block of seven aptitudinal predictors consisting of SES, IQ and sex main effects and the four possible interactions between these three main effects. The fourth column contains results for I_tA_f, a block of the seven aptitude-treatment interactions possible between curricular treatment and SES, IQ or sex. As well, the bottom row of Table 1 displays, for I_t, A_f and I_tA_f, the median percentage of unique variance accounted for across the 17 learning outcomes.

In performing the multiple regression analyses reported in Table 1, use was made of *Method II* of Overall and Spiegel (1969), described in more detail by Fraser, McArthur and West

TABLE 1

F values for 17 learning outcomes using I_i (instruction), A_j (aptitudes) and I_iA_j interactions as predictors

Predictor / Criterion	I_i	Predictors in blocks		Significant individual predictors passing blocking criteria			
		A_j	I_iA_j	Main A_j effects SES	IQ	Sex	A_j and I_iA_j interactions
TOES 1 (Library)	2.0	9.0****	1.1		H 59.0****		
2 (Index)	0.0	5.4****	0.2		H 30.3****	F 4.3*	
3 (Scales)	1.3	7.6****	0.5		H 50.5****		
4 (Averages)	1.6	6.3****	1.5		H 39.4****	M5.6*	4.5* (I*sex) 3.8* (I*sex*SES)
5 (Charts)	0.0	8.9****	0.2		H 57.5****		
6 (Graphs)	0.1	9.9****	0.5		H 64.2****		
7 (Compr)	0.0	8.4****	0.5		H 56.5****		
8 (Exptl)	0.4	11.7****	0.7	H 3.9*	H 73.0****		
9 (Concl)	2.9	7.6****	0.2		H 50.1****		
TOUS P (Philos)	0.6	6.8****	0.6	H 5.0*	H 39.9****		
H (Hist-soc)	0.1	15.3****	0.4		H 98.7****		
N (Normal)	0.0	5.8****	0.4		H 32.7****	F 6.9**	
Att. I (Interest)	0.2	0.9	0.8				
E (Enjoy)	A 13.7***	0.8	1.5				3.8* (I*SES)
S (Soc Imp)	2.4	0.6	0.9				
Enquiry	0.2	0.9	0.5				
Scientific atts.	0.1	4.7****	0.3	H 4.8*	H 20.3****	M7.8**	
Critical F ($\alpha = .05$)	$F_{1,326} = 3.8$	$F_{7,326} = 2.0$		$F_{1,326} = 3.8$			
Median R^2(%)	0.0	8.0	0.7				

* $p < .05$, ** $p < .01$, *** $p < .001$, **** $p < .0001$
A ASEP, H high, M males, F females

N = 343 class aptitude cells (46 schools)

(1975). For each analysis, the full regression equation contained 16 predictor variables consisting of the appropriate pre-test, curricular treatment, SES, IQ, sex and the 11 interactions possible between treatment, SES, IQ and sex. The technique employed, therefore, would have as its analogue a four-way analysis of covariance with pre-test as covariate and with curriculum, SES, IQ and sex as main effects. The multiple regression analysis used, however, had three advantages over its analysis of covariance analogue. Firstly, it permitted the use of IQ as a continuous predictor variable without any need for categorization; secondly, it allowed a least squares solution to the disproportionate cell data that existed; and, thirdly, it readily permitted the use of the blocks of predictors, A_j and I_iA_j, represented in Walberg's model.

The results in Table 1 indicate that curricular treatment I_i was a relatively weak predictor of learning, with only one F value significant at the 0.05 level and a median of 0.0 percent of unique variance accounted for. The significant difference, which was in favour of the ASEP group, occurred for Attitude E (Enjoyment of Science Lessons). The aptitude block A_j proved to be a potent predictor of learning, with 13 of the F values in Table 1 significant and with a median of 8.0 percent of unique variance explained. For the block of I_iA_j treatment-aptitude interactions, no significant F values were found and the median amount of unique variance accounted for was only 0.7 percent.

The last four columns in Table 1 contain F ratios for the individual terms which comprised the A_j and I_iA_j blocks. For these individual terms, significance tests were only carried out if a blocking criterion was passed. For individual A_j main effects and interactions, significance tests were carried out only when the A_j block had yielded significance at the 0.1 level. Individual I_iA_j interactions were tested for significance only when either I_i, the A_j block or the I_iA_j block had yielded significance at the 0.1 level. The purpose for applying these blocking criteria was to minimize the number of spuriously significant findings that could arise if all possible 238 individual tests were carried out. In Table 1, entries have been made in the last four columns only for those predictors which passed the blocking criteria and which were individually significant at the 0.05 level.

The results for individual A_j and I_iA_j predictors in Table 1 show that IQ was the most potent predictor of learning. It is interesting to note, too, that IQ was a significant predictor of all twelve cognitive learning outcomes but of only one of the five attitudinal outcomes. Pupils of high SES performed significantly better than those of low SES on TOES Skill 8 (Design of Experimental Procedures), TOUS Scale P (Philosophical aspects) and the Scientific Attitudes scale. Boys were significantly better than girls for TOES Skill 4 (Averages, Percentages and Proportions) and the Scientific Attitudes scale while girls performed significantly better than boys on TOES Skill 2 (Index and Table of Contents) and TOUS Scale N (Normality of Scientists).

The findings for TOES Skill 4 (Averages, Percentages and Proportions) are of particular interest. Firstly, boys performed significantly better than girls. Secondly, while the curriculum main effect was small, there was a significant curriculum by sex interaction. An examination of scores adjusted for all other terms in the regression model revealed that ASEP boys and control boys had a similar performance. On the other hand, adjusted scores indicated that girls in the control group achieved much better on TOES Skill 4 than girls in the ASEP group.

ASEP and enjoyment of science lessons. Earlier, it was seen that the only significant difference found between ASEP and control groups was for Attitude E (Enjoyment of Science Lessons). As well, a significant curriculum by SES interaction occurred for Attitude E. In Figure 3, raw score means have been plotted to illustrate these findings.

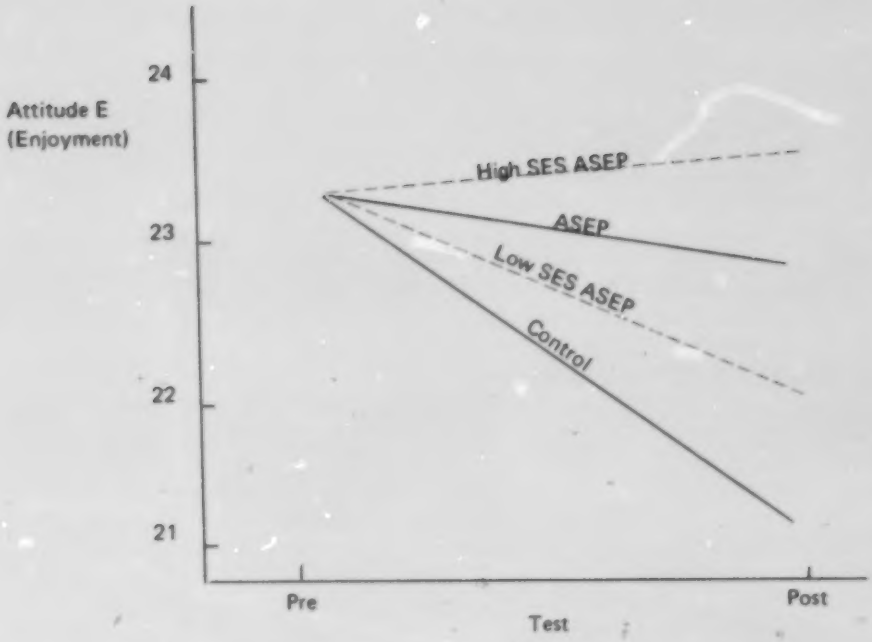


FIGURE 3: Raw score changes in enjoyment

Firstly, Figure 3 shows that while ASEP and control pupils expressed similar enjoyment of science lessons near the beginning of the year, both groups declined in enjoyment throughout the year. However, while the control group declined 2.1 raw score points, the ASEP group declined only 0.4 points. That is, in the ASEP group, there was an arrest of 1.7 points, or about one third of a standard deviation, in the decline of enjoyment experienced by the control group. For the Attitude E criterion, curricular treatment accounted for 2.7 percent of unique variance.

Secondly, Figure 3 illustrates an interesting interaction effect between curricular treatment and SES for Attitude E. For the control group, examination of adjusted scores revealed that performance on Attitude E was independent of SES; for this reason, Figure 3 contains only the one line for the control group. On the other hand, adjusted scores for the ASEP group showed clear superiority of high SES pupils over low SES pupils. The dotted lines in Figure 3 indicate that high SES ASEP pupils experienced a slight increase in enjoyment of 0.2 points throughout the year while the low SES group experienced a decline of 1.1 points.

The present findings related to enjoyment are similar to those reported in other research. Walberg, House and Steele (1973) reported a decline in American pupils' general attitude to school as they advanced through the grade levels. Mackay (1970) and Gardner (1973) found a decrease in Australian pupils' enjoyment of learning physics as they proceeded through the P.S.S.C. course. Choppin (1974) found that British Nuffield science pupils experienced a decline in liking for science between Form 3 and Form 4, but that the Nuffield pupils' liking for science was greater than that of non-Nuffield students. Furthermore, the present finding that curricular

treatment was a relatively weak effect by comparison with other variables has been endorsed by Gardner (1975a, p. 29), Walberg (1969, p. 186) and Choppin (1974, p. 202).

The impact of ASEP on classroom climate. It is possible that ASEP may differ from traditional programs, not only in its impact on learning outcomes, but also on its impact in creating a different and more favourable classroom climate. In order to explore this possibility in the present study, a set of nine classroom climate dimensions was administered as criterion measures at the end of the second school term.

These nine classroom climate dimensions, listed in Table 2, were based on the Learning Environment Inventory (LEI) scales of Walberg and Anderson. The first eight dimensions in Table 2 were selected from the larger set of LEI scales (Anderson, 1971) because of their relevance to the ASEP situation. The actual items used were also based on the original LEI scales but contained numerous modifications to enhance suitability and readability for Australian pupils in their first year at secondary school. The ninth scale, Individuality, was developed for the purposes of the present study in an attempt to measure an aspect of classroom climate which, although of great importance in the ASEP situation, was not covered by the original LEI battery. For these nine classroom climate measures, the Cronbach α reliability ranged from 0.45 to 0.80 with a median of 0.63.

TABLE 2

F values for 9 classroom climate dimensions using I_j (instruction), A_j (aptitudes) and I_jA_j interactions as predictors

Predictor / Criterion	I_j	Predictors in blocks		Significant individual predictors passing blocking criteria			
		A_j	I_jA_j	Main A_j effects SES	IQ	Sex	A_j and I_jA_j interactions
Diversity (+)	A 0.2	1.2	1.3		H 5.9*		
Speed (-)	A 1.6	2.0*	1.1		H 9.6*		4.0* (IQ*sex)
Environment (+)	A 3.8*	0.7	0.9				4.6* (I*SES*IQ)
Goal Direction (+)	A 0.2	1.3	0.1				
Satisfaction (+)	A 7.5**	0.9	0.7				
Disorganization (-)	A 0.2	1.4	0.2				
Difficulty (-)	A 3.0	3.3**	0.7		H 17.3***		
Competition (-)	A 2.9	1.7(1)	0.5		H 7.4**		
Individuality (+)	A 15.3***	2.0*	0.7		H 8.4**		
Critical F (α= .05)	$F_{1,137}$=3.8	$F_{7,137}$=2.0		$F_{1,137}$=3.8			
Median R^2(%)	1.8	6.3	2.9				

* p = .05, ** p = .01, *** p = .001, **** p = .0001
A ASEP mean in favourable direction, H High IQ mean in favourable direction
+ high score is favourable, - low score is favourable
N = 153 class aptitude cells (20 schools)

The sample used in this part of the study was smaller than that used previously, consisting of 10 of the original ASEP classes and 10 of the original control classes. Class aptitude cell

sampling units were again used to minimize confounding between A_j and E_k. The fact that this group of ASEP pupils was not significantly different from the control group on any of the 17 learning outcome pre-tests nor on SES, IQ and sex was again taken as evidence that little confounding between I_i and A_j would be present. The possibility of confounding between I_i and E_k could not be dismissed, however, especially in the absence of pre-test climate measures. That is, there existed the possibility that observed differences between the ASEP and control group might be attributable in part to the fact that the teachers and schools choosing to use ASEP were different in some way from those choosing to use conventional materials. Nevertheless, information from a questionnaire (Northfield, private communication) was analyzed to show that the ASEP and control teachers used in the classroom climate study were not significantly different on four dimensions, namely, sex, years of teaching experience, attitude to structure in science lessons and attitude to pupil individuality.

Table 2 provides the results for classroom climate using a fifteen-term regression model containing curriculum I_i, the A_j block of seven aptitudinal variables and the I_iA_j block of seven aptitude-treatment interactions. These results indicate significant differences in favour of the ASEP group on the Environment, Satisfaction and Individuality scales. Furthermore, while differerences on the remaining scales were nonsignificant, the direction of the differences favoured the ASEP group in all cases. That is, by comparison with control students, ASEP pupils rated their classes as being characterized by more diversity, less speed, a better environment, more goal direction, more satisfaction, less disorganization, less difficulty, less competition and more individuality. This pattern of findings is both striking and consistent with the earlier finding that ASEP pupils performed significantly better than control students on the Enjoyment of Science Lessons scale. Together, Table 1 and Table 2 suggest that, while ASEP's impact on learning outcomes was relatively small, the impact on classroom climate was appreciable.

Table 2 also shows that aptitude, particularly IQ, was a potent predictor of pupils' perception of classroom climate. Pupils of high IQ perceived their classes as significantly less speeded, less difficult, less competitive and more individual than did pupils of low IQ. As well, an interesting significant interaction of IQ with sex was detected for the Speed scale. While boys of high IQ perceived their classes as considerably more speeded than boys of low IQ, girls' perceptions of speed were independent of their IQ.

Some caution, however, is needed in interpreting the findings related to climate. Not only was the sample considerably smaller than that used for learning outcomes, but confounding between I_i and E_k must be considered a possibility, especially in the absence of climate pre-tests. Nevertheless, the pattern of findings in which the direction of differences favoured ASEP students on all nine climate dimensions, suggests that the impact of ASEP on classroom climate was relatively large.

Conclusion

Consideration has been given to an evaluation of ASEP conceptualised in terms of Walberg's model, which contains instructional, aptitudinal and environmental variables. Multiple regression analyses, which employed an appropriate sampling unit in order to minimize confounding due to environment, revealed that instruction was a relatively weak predictor of learning by comparison with aptitude. Out of the 17 learning outcomes considered in the study, ASEP classes were found to be significantly superior to control on one criterion, enjoyment of science lessons. While both ASEP and control pupils declined in enjoyment during the year, the decline

in the ASEP group was substantially smaller than in the control group. Nevertheless, while the impact of ASEP on learning outcomes had been relatively small, a second set of analyses revealed that ASEP's impact in promoting a favourable classroom climate was appreciable.

References

ANDERSON, G.J. *The assessment of learning environments: A manual for the Learning Environment Inventory and the My Class Inventory.* Halifax: Atlantic Institute of Education, 1971.

AUSTRALIAN SCIENCE EDUCATION PROJECT. *Unit 401: A guide to ASEP,* Melbourne, 1974.

BARKER, R.G. *Ecological Psychology: Concepts and Methods for Studying the Environment of Human Behavior.* Stanford University Press, Stanford, 1968.

BRACHT, G.H. & GLASS, G.V. "The external validity of experiments." *American Educational Research Journal,* **5**: 437-474, 1968.

BRUNSWIK, E. *Perception and the Representative Design of Psychological Experiments.* University of California Press, Berkeley, 1956.

CAMPBELL, D.T. & STANLEY, J.C. "Experimental and quasi-experimental designs for research on teaching", in N.L. Gage (Ed.), *Handbook of Research on Teaching.* Rand McNally, Chicago, 171-246. 1963.

CATTELL, R.B. "Psychological theory and scientific method," in R.B. Cattell (Ed.), *Handbook of Multivariate Experimental Psychology.* Rand McNally, 1-18, 1966.

CHOPPIN, B.H. "The introduction of new science curricula in England and Wales." *Comparative Education Review,* **18**: 196-206, 1974.

COLEMAN, J.S. *Equality of Educational Opportunity.* U.S., Department of Health, Education and Welfare, 1966.

CONGALTON, A.A. *Status and Prestige in Australia.* Cheshire, Melbourne, 1969.

CRONBACH, L.J. "The two disciplines of scientific psychology." *American Psychologist,* **12**: 671-684, 1957.

FRASER, B.J. "Selecting evaluation instruments." *Research in Science Education,* **4**: 99-111, 1974.

FRASER, B.J. & FISHER, D.L. "Changes in pupils during a year of ASEP in Tasmania." *Australian Science Teachers Journal,* **20** (*2*): 89-96, 1974.

FRASER, B.J., McARTHUR, J.T. & WEST, L.H.T. "Some issues in the simulation of ANOVA by multiple regression." This volume, 1975.

GARDNER, P.L. "Changes in attitudes of PSSC physics students: a second look." *Australian Science Teachers Journal,* **19** (*1*): 71-78, 1973.

GARDNER, P.L. "Attitudes to science: a review." *Studies in Science Education,* **2**: 1-41, 1975(a).

GARDNER, P.L. "Attitudes of PSSC physics students: relationships with personality and teacher behaviour." *Australian Science Teachers Journal,* **21** (*1*): 75-85, 1975(b).

LINDQUIST, E.F. *Design and Analysis of Experiments in Psychology and Education.* Houghton Mifflin, Boston, 1953

LORD, F.M. "Elementary models for measuring change", in C.W. Harris (ed.), *Problems in Measuring Change.* University of Wisconsin Press, Madison, 21-38, 1963.

MACKAY, L.D. "Changes in Victorian physics students during two years of physics study." *The Australian Physicist*, **7**: 103-109, 1970.

OVERALL, J.E. & SPIEGEL, D.K. "Concerning least squares analysis of experimental data." *Psychological Bulletin*, **72**: 311-322, 1969.

PECKHAM, P.D., GLASS, G.V & HOPKINS, K.D. "The experimental unit in statistical analysis." *Journal of Special Education*, **3**: 337-349, 1969.

PEREBOOM, A.C. "Some fundamental problems in experimental psychology: an overview." *Psychological Reports*, **28**: 439-455, 1971.

SNOW, R.E. "Representative and quasi-representative designs for research on teaching." *Review of Educational Research*, **44**: 265-291, 1974.

WALBERG, H.J. "A model for research on instruction." *School Review*, **78**: 185-200, 1970.

WALBERG, H.J. "Models for optimizing and individualizing school learning." *Interchange*, **2**: 15-27, 1971.

WALBERG, H.J., HOUSE, E.R. & STEELE, J.M. "Grade level, cognition and affect: a cross section of classroom perceptions." *Journal of Educational Psychology*, **64**: 142-146, 1973.

EXPERIMENTAL STUDIES OF TEACHER STRUCTURING BEHAVIOUR IN ASEP CLASSROOMS

Colin N. Power & Richard P. Tisher

Introduction

Whatever curriculum materials the science teacher has elected to use, he must create a learning environment in which new understandings and sensitivities can develop. His initial comments and subsequent moves he makes to a large degree steer the activities of students, focussing their attention on certain goals, content and procedures as opposed to others. Such comments and moves, called structuring moves by Bellack *et al.* (1966), represent pivotal points in classroom discourse and serve to initiate, steer and terminate classroom activities. Similarly, Rosenshine (1971) uses the term structuring to refer to teacher comments at the beginning or end of a lesson, and to teacher comments before and after he asks a question.

The term "structuring" implies that a major function of such moves is to provide a "structure" or scaffolding for ideas and specifics to come, or for those which have just been presented. The concept of structure features prominently in the writings of many curriculum and learning theorists. For example, Ausubel (1968) uses the term "organizer" to describe the class of structuring moves designed to present the "relevant subsuming concepts" at a high level of generality. He argues strongly for the inclusion of such organizers in introductory overviews to provide "anchoring foci for the reception of new material". Organizers presented beforehand are commonly called "advance organizers" and those contained in structuring summaries after presentation are called "post organizers". The latter are intended to consolidate and integrate previously established ideas.

In general, the data do indicate that advance organizers do facilitate learning from meaningful lesson material (Ausubel & Fitzgerald, 1961, 1962; Ausubel & Youssef, 1963; Fitzgerald & Ausubel, 1963). In most cases the provision of advance organizers did improve learning and retention; however, in no case has the improvement been very large. In a study using written materials, Bauman and Glass (1969) found that post-organizers were somewhat more effective than pre-organizers. In a detailed naturalistic study, Wright (1969) found that the extent to which teachers use structuring to integrate learning experiences at the end of a lesson (i.e. post-organizers) was more strongly associated with achievement than teachers' use of recapitulation at the beginning of the next lesson. Studies using quantitative techniques developed by Anderson (1971) and Carss (1973) to analyse the structure of teacher-student interaction and curriculum materials, indicate that the overall structure and organization of lessons and curriculum materials is important in science teaching. In these studies (Anderson, 1967; Butterworth, 1974; Browne & Anderson, 1975; Clarke, 1973, 1975), students who used more carefully integrated and organized materials or who experienced more structured lessons learned more, and expressed greater satisfaction, than students in less structured modes.

When one begins to look a little more carefully at the data, the relationship between structuring and achievement is less clear cut. In a reanalysis of the Bellack *et al.* (1966) data, Furst (1967) found that superior teachers made moderate use of structuring. Other studies also seem to indicate that the optimal level of structuring may depend on the properties of the student, the task, and the teacher. For instance, dependent-prone and culturally deprived students seem to prefer and learn more in structured situations, whereas independent, very able students

and college males do not (Brophy & Evertson, 1974, Grimes & Allinsmith, 1961; McKeachie, Lin & Mann, 1971).

It would appear that for many students carefully organized and structured learning experiences and materials, and advance and post organizers are likely to facilitate the acquisition and retention of content.

The somewhat traditional values implicit in Ausubel's position appeal to many science educators who find in his writings a theoretical rationale for the notion that the chief responsibility of teachers and curriculum materials is to present an organized body of knowledge to students. But other theorists (e.g. Rogers, 1969, Suchman 1960), curriculum workers (e.g. Samples, 1971), and teachers less concerned with mastery of content, believe that highly structured curricula and teaching are dehumanizing, get in the way of student thinking and inhibit the development of inquiry skills.

This rift in value orientation soon became apparent to us in discussions held with ASEP staff and trial teachers in the early phases of our research project. Whereas some ASEP staff believed in and produced highly structured units, others argued for and developed units which provided students with relatively little explicit guidance as to how to tackle problems. Furthermore, in the more open-ended units, teachers are requested *not* to advise students how to plan their work, but merely to introduced problems and explain to students what is required of them. How much structure and guidance to provide in lessons also represented a concern frequently expressed by trials teachers. The critical question becomes: *How much structure should be provided within the materials and by the teacher, and at what points in the learning sequence?* It is to this question that the experimental studies described below address themselves. The studies also attempt to take into account the value orientation of teachers, since teachers differ in the extent to which they believe that the learning environment should be open rather than highly structured.

Two studies were carried out using different units ("Light Forms Images" and "How Many People") and different samples. In the first study, we were largely concerned with the effects of variations in the amount and source of structuring. In the second, the concern was with the temporal locus of structuring.

Study I: Amount and source of structure

Procedures used in the selection of the sample of experimental teachers included in the first study have been described elsewhere (Tisher & Power, 1973) as has the study itself (Tisher & Power, 1975). Accordingly the description provided here will be brief. Fifteen Diploma in Education students which included three representatives of each of five value orientations (ranging from open to traditional) were asked to teach the core of the ASEP unit "Light Forms Images" to classes of ten to twelve randomly selected Grade 9 pupils. Three of the five lessons required for the core material were videorecorded and analysed using the SABIC coding scheme (Tisher & Power, 1973). Criterion-referenced achievement and attitude pre-tests were administered before teaching on the unit began, and parallel post-tests administered during the sixth lesson.

The classes of the 15 Diploma students were allocated to three different treatments so that a student teacher from each value cluster was assigned to one of the treatments. The treatments were as follows:

Treatment 1 : Additional structuring. The five classes in this treatment used the ASEP pupils' guide, record book and ancilliary equipment. At the beginning and end of each lesson the teachers led a class discussion on the salient points which had been covered previously and indicated those points which would arise in the ensuing lesson: that is, they involved the whole class in review and overview at the beginning and end of each lesson.

Treatment 2 : ASEP materials. The classes in this treatment used the ASEP pupils' guide, record book and ancilliary equipment. Their activities were guided by the instructions in the curriculum materials and the teacher acted as a resource person.

Treatment 3 : Unstructured. In this treatment the classes were supplied with all the equipment and the aids necessary to study the concepts in the ASEP unit "Light Forms Images" but they were encouraged to explore on their own with the teacher acting as a resource person and, where appropriate, guide.

All diploma students were trained to use ASEP materials and to follow the treatment to which they were assigned. Those assigned to treatment 3 were familiar with the unit "Light Forms Images" and sought the same objectives.

Value orientations were as follows:

I Persons valuing open and free communication among students and teacher; an enquiry-oriented, pupil-centred rather than teacher-dominated classroom.

II as above – except that they do not believe students should devise their own experiments or be free to communicate among themselves.

III as above – except less value placed on adequate teacher-pupil communication.

IV as above – except that greater value is placed on rules to guide student behaviour.

V Persons who believe in a teacher dominated classroom, impersonal relationships, whole class teaching, rules and reliance on curriculum materials.

Value orientation I is most congruent with ASEP; V is most dissonant.

The data were analysed using an unweighted-means analysis of covariance technique with pre-test scores as the covariate. Treatments and teacher values were assumed to be fixed effects. The results are summarized in Tables 1 and 2. The unit of analysis is the individual pupil – pupils were allocated at random to classes and in all classes pupils had the opportunity to respond individually and independently to the self-paced materials.

The results are complex and only the essence of our interpretation is reproduced here since a more detailed discussion appears elsewhere (Tisher & Power, 1975). It seems that so far as student achievement is concerned, student performance under treatment 1 (Additional Structuring) is superior to treatment 2 (ASEP materials) and that, on the average, treatments 1 and 2 are superior to 3 (Unstructured). It seems that a moderate amount of teacher structuring, especially by persons whose educational values are not completely congruent with the ASEP approach, does facilitate learning. On the other hand, pupils of teachers whose values are more open and more congruent with those of ASEP may perform less well on criterion-referenced tests if the teacher diverges considerably from the printed curriculum guide.

With respect to changes in student attitudes, it seems that students may appreciate additional structuring by the teacher when the teachers values are more congruent with those of ASEP (open, pupil centered, inquiry oriented). On the other hand, students taught by student teachers with a more traditional orientation toward teaching enjoyed the unstructured situation more than the structured ones. A balancing mechanism may be operating.

TABLE 1
Analysis of Covariance : Attitude

Source of Variation	Sum of Squares	Df	Mean Square	F	P
Treatment	64659	2	32329	2.62	*
Values (B)	224390	4	56097	4.55	***
AXB	219726	8	27465	2.23	**
Error	1653883	134	12342		

* significant at 10% level
** significant at 5% level
*** significant at 1% level

Attitude Adjusted Cell Means

Treatment	Values				
	I	II	III	IV	V
1. Additional Structuring	4.61	49.72	-13.77	25.90	-74.67
2. ASEP materials	23.74	-53.38	-1.02	54.67	-79.44
3. Unstructured	-45.39	-3.14	28.80	163.18	45.13

Note: Using M2.1 (-23.74) as a reference, significant differences are: –
M2.1 – M3.IV at 1% level
M2.1 – M1.V at 10% level
M2.1 – M2.V at 5% level

TABLE 2
Analysis of Covariance : Achievement

Source of Variation	Sum of Squares	Df	Mean Square	F	P
Treatment (A)	62.0	2	31.0	5.29	***
Values (B)	121.6	4	30.4	5.18	***
AXB	105.6	8	13.2	2.25	**
Error	785.8	134	5.9		

* significant at 10% level
** significant at 5% level
*** significant at 1% level

Achievement Adjusted Cell Means

Treatment	Values				
	I	II	III	IV	V
1. Additional Structuring	10.8	11.66	9.71	7.85	11.49
2. ASEP materials	9.25	11.10	9.43	8.91	8.90
3. Unstructured	6.14	10.64	9.35	8.56	8.70

Note: Using M2.1 (-9.25) as a reference, significant differences are: –
M2.1 – M3.I at 1% level
M2.1 – M1.II at 5% level
M2.1 – M2.II at 10% level
M2.1 – M1.V at 5% level

Study II : Temporal locus of structuring moves

Just as the study described above formed part of a larger project funded by the AARCDE, so too did the second experimental study. This study involved 20 teachers selected by the procedures used in the study above, that is, selected so that teachers representing each of several value orientations towards science teaching were included in the sample.

In the experimental phase of this project, twelve science Diploma of Education students were asked to teach the core of the ASEP unit "How Many People". Four representatives of each of the following value orientations were included in the sample.

I Persons who value adequate communication among students and teachers, supportive personal relationships, student inquiry and first hand experience (i.e. ASEP congruent).

II Persons who value adequate communication, supportive relationships, but who believe the teacher should be the dominant actor and place less emphasis on student activity (i.e. mixed mode).

III Persons who believe the teacher should be the dominant actor, that there should be restrictions on student initiative, communication with students and personal relationships (ASEP dissonant).

Classes of the student teachers comprised 10 to 12 randomly selected Grade 9 pupils. Two of the six lessons required for the core material were videorecorded and the video-tapes were analysed using the SABIC coding system (Tisher & Power, 1973). Achievement and attitude pre- and post-tests were administered. Pre-test measures used as covariates in this study were:

1. Inter D Scholastic Aptitude
2. Understanding of ecological concepts
3. Understanding of population ecology
4. Attitudes towards pollution issues
5. Attitudes towards population control
6. Attitudes towards resource utilization
7. Interest in science lessons

Post-test measures used included:

8. Understanding of ecological concepts
9. Understanding of population ecology
10. Attitudes towards pollution issues
11. Attitudes towards population control
12. Attitudes towards resource utilization
13. Interest in ASEP science lessons
14. Perceived amount of learning
15. Satisfaction with organizational context.

Classes were assigned at random to one of four treatments:

Treatment 1 : ASEP materials. The classes in this treatment used the ASEP pupils' guide, record book and ancillary equipment. Their activities were guided by the instructions in the curriculum materials and the teacher acted as a resource person.

Treatment 2 : Pre-structuring. The classes in this treatment used the material above except that in the first lesson an overview highlighting the basic ideas developed in the unit was outlined by the teacher. The BSCS film "Population Ecology" was also shown in this lesson to assist in the development of subsuming concepts (i.e. advance organizer).

Treatment 3 : Mid-structuring. As for 2 – except that structuring occurred in the middle lesson.

Treatment 4 : Post-structuring. As for 2 – except that structuring was provided in the last lesson.

All Diploma students undertook a training course designed to assist them to use the unit "How Many People". In the case of Diploma students assigned to structuring treatments, additional training and discussion relevant to the particular unit used was undertaken. Observational data indicate that the Diploma students did in fact adopt the roles and strategies for which they were trained.

The data were analysed using multivariate analysis of variance with the set of pre-tests described earlier as covariates and the set of post-tests as the criteria. Values and Treatments were assumed to be fixed effects. A summary analysis of variance table is presented in Table 3.

TABLE 3

Study II, Summary Analysis of Variance Table

Source	df	Multivariate F			Stepdown F	
		F	df	sig level	$v_{13}/v_{8\text{-}12}$	$v_{14}/v_{8\text{-}13}$
Const (M) & covariates (X)	7	-	-	-	-	-
Values (A) elim M,X	2	1.13	(16,210)	NS.		
Treatments (B) elim M,X,A	3	2.23	(24,305)	***	9.86***	1.94
Interaction elim M,X,A,B	6	1.76	(48,521)	***	5.76***	2.10*
Within	112					
Total	130					

Univariate F-ratios

Variable	Treatment	Values	Interaction
8	1.42	1.51	.94
9	.52	.59	.65
10	1.30	.25	1.02
11	1.01	2.37	1.68
12	.80	.08	1.31
13	12.03***	2.78	5.63***
14	3.83**	3.01*	6.75***
15	1.40	2.21	1.89

*** p < .001, ** p < .01, * p < .05

Data analysis revealed a significant interaction effect. Examining the stepdown results, we find that the interaction effect stems largely from variable 13 and, to a lesser extent, from variable 14. Univariate analyses of variance confirm that the adjusted criterion variables most affected by variations among treatment-value cells were variables 13 and 14, which measured student interest in the ASEP lessons they experienced and their perception of the amount they learned, relative to other science lessons.

Referring to the estimated means for interest in ASEP lessons (Figure 1) we see that, in general, students in the experimental groups experiencing some additional structuring by film and teacher (pre, mid and post) expressed a higher level of interest and satisfaction, particularly groups experiencing post structuring. The trend in the data for perceived amount of learning followed the same pattern.

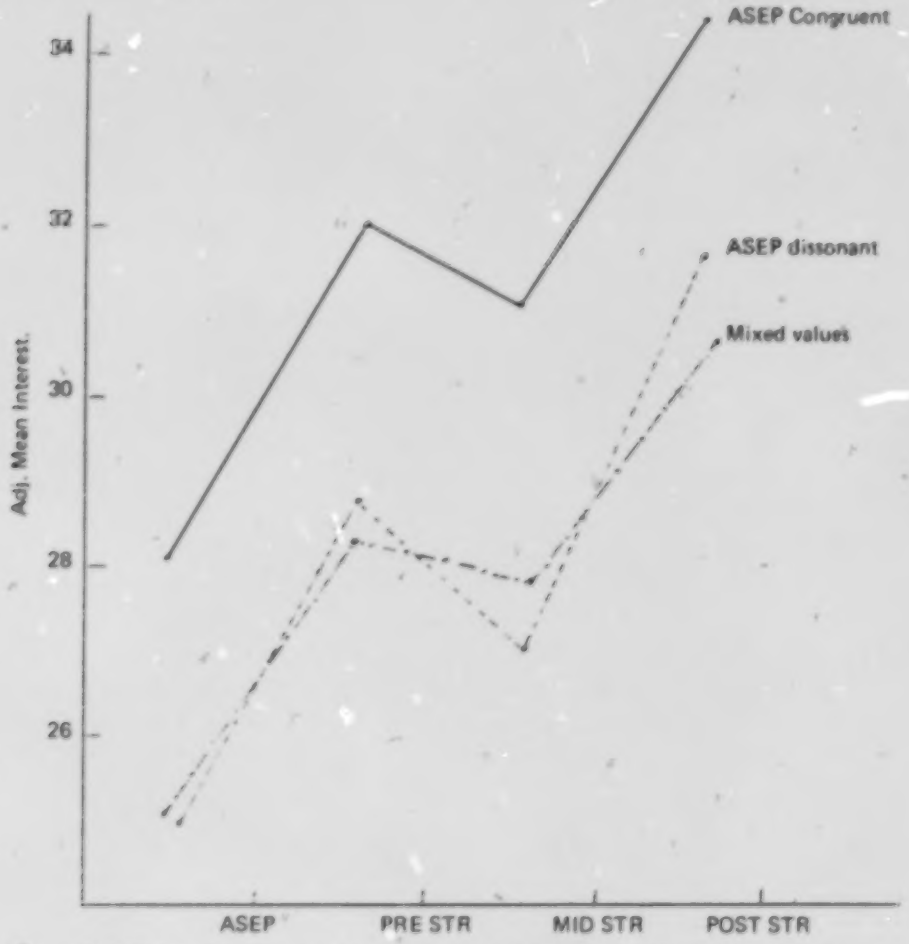


FIGURE 1: Treatments x Values Interaction II Estimated Means after Covariance Adjustment on Student Interest in ASEP Lessons.

Eliminating the interaction affects, the main effect for treatments was found to be significant ($p < .001$). Estimates of the effects for variables 13 and 14 reveal a clear trend: post-structuring > pre-structuring > mid-structuring > ASEP materials. Contrasts revealed that post-structuring was superior to pre-structuring, supporting the Bauman & Glass and the Wright results. Univariate analyses revealed a significant main effect for variables 13 and 14 only, and the same general trends. While multivariate F revealed no significant main effects for values, one univariate F (variable 14) approached significance. The trend in the data consistently favoured classes taught by ASEP-congruent teachers, but the lack of significant differences indicates this effect is small. It is interesting to note that whereas treatment and the interaction effects were significant, the effects were largely on student interest and perception of learning. The treatments and general educational values of teachers did not appear to affect measured student learning or their attitudes towards ecological issues.

Conluding comment

When interpreting the results of these two studies it should be borne in mind that in both cases this was both the teachers' and the students' initial exposure to ASEP and that this exposure was brief. The effects revealed may or may not persist in time. Furthermore, the extent to which the interaction effects obtained are generalisable is also open to question in view of the small sample of teachers involved and their lack of experience.

Given these cautions, one is tempted to suggest that in the initial stages of using ASEP materials the teacher needs to take time out towards the end of the unit to assist students to integrate and consolidate the experiences which the student has had.

In individualised programs like ASEP and ISCS the teacher is called upon not to teach, but to create a viable learning environment. Rather than playing the role of a broadcasting station, the teacher must devote more time and thought to ways of stimulating meaningful learning and interest; and to ways of integrating the diverse experiences of individuals so that clear, stable and usable concepts and generalisations will emerge. As conductor of the learning orchestra, the teacher must be responsible for the subtle blending of the notes from many instruments which produces a work of art rather than chaos. In this role, our results indicate that there must be a score: a grand plan which is sufficiently clear to give structure, meaning and purpose to activity, yet sufficiently flexible to allow for individual talent and artistry.

References

ANDERSON, O.R. *Quantative Analyses of Structure in Teaching.* Teachers College Press, New York, 1971.

ANDERSON, O.R. "Research on structure in teaching." *Journal of Research in Science Teaching,* **3**: 219-230, 1974.

AUSUBEL, D.P. *Educational Psychology: A Cognitive View.* Holt, Rinehart & Winston, New York, 1968.

AUSUBEL, D.P. & FITZGERALD, D. "The role of discriminability in meaningful verbal learning and retention." *Journal of Educational Psychology,* **52**: 266-274, 1961.

AUSUBEL, D.P. & FITZGERALD, D. "Organizer, general background and antecedent learning variables in sequential verbal learning." *Journal of Educational Psychology,* **53**: 243-249, 1962.

AUSUBEL, D.P. & YOUSSEF, M. "The role of discriminability in meaningful parallel learning." *Journal of Educational Psychology*, **54**: 331-336, 1963.

BAUMAN, D.J. & GLASS, G.V. "The effect on learning of the position of an organizer." Paper read at annual meeting of the American Educational Research Association, Los Angeles, 1969.

BELLACK, A.A., KLIEBARD, H.M., HYMAN, R.T. & SMITH, F.L. *The Language of the Classroom.* Teachers College Press, Columbia, New York, 1966.

BROPHY, J. & EVERTSON, C.M. "Process-product correlations in the Texas Teacher Effectiveness Study: A final report." Report No. 74-4. University of Texas at Austin, June, 1974.

BROWN, R. & ANDERSON, O.R. "Pupil perception of and achievement in secondary school biology lessons as related to lesson Kinetic Structure Analysis." *Journal of Educational Psychology* (in press).

BUTTERWORTH, T. "A study of the effect of lesson verbal structure as aspects of student effective learning in freshman college biology." Cited in O.R. Anderson, "Research on structure in Teaching." *Journal of Research in Science Teaching*, **11**: 219-230, 1974.

CARSS, B. "Content analysis as a technique for determining the structure of materials." *Science Education Research*, (ASERA), 107-117, 1973.

CLARK, J. "The role of the content and structure of curriculum material in recognition." *Science Education Research*, (ASERA), 119-141, 1973.

CLARK, J. "Thematic analysis of science lessons." *Classroom Interaction Newsletter.* (in press).

FITZGERALD, D. & AUSUBEL, D.P. "Cognitive versus affective factors in the learning of controversial material." *Journal of Educational Psychology*, **54**: 73-78, 1963.

FURST, N.F. "The multiple languages of the classroom." Paper read at the American Educational Research Association Annual Meeting, New York, 1967.

GRIMES, J.W. & ALLINSMITH, W. "Compulsivity, anxiety and school achievement. *Merrill-Palmer Quarterly*, **7**: 247-269, 1961.

McKEACHIE, W.J., LIN, Y. & MANN, W. "Student ratings of teacher effectiveness: Validity studies." *American Educational Research Journal*, **8**: 435-445, 1971.

ROGERS, C.R. *Freedom to Learn.* Charles Merrill, Columbus, Ohio, 1969.

ROSENSHINE, B. "Teaching behaviour related to pupil achievement: a review of research." In I. Westbury & A. Bellack (ed). *Research into Classroom Processes.* Teachers College Press, New York, 1971.

SAMPLES, R. *Growing: Environmental Studies Project.* American Geological Institute, Boulder, Colorado, 1971.

SUCHMAN, T.R. "Inquiry training in the elementary school." *The Science Teacher*, **27**: 42-47, 1960.

TISHER, R.P. & POWER, C.N. "The effects of teaching strategies in mini-teaching and micro-teaching situations where Australian Science Education Project material are used." Mimeo report of AACRDE project, University of Queensland, 1973.

TISHER, R.P. & POWER, C.N. "A study of the effects of teaching strategies in A.S.E.P. classrooms." *Australian Journal of Education*, (in press).

TATSUOKA, M. "Multivariate analysis in educational research." In F. Kerlinger (ed). *Review of Research in Education 1.* F.E. Peacock, Itasca, Illinois, 1973.

WRIGHT, C.J. & NUTTHALL, G. "Relationships between teacher behaviors and pupil achievement in three experimental elementary science lessons." *American Educational Research Journal*, **7**: 477-490, 1970.

CHANGING PERSPECTIVES OF THE SCIENCE TEACHER EDUCATION THAT MAY BE ASSOCIATED WITH CURRICULUM DEVELOPMENT

Jeff R. Northfield

Introduction

> "It is now time to call a halt to writing syllabuses or developing new curriculum materials until there are effective teacher education and evaluation programmes to accompany them." (Ramsey, 1972)

This statement provides one example of the growing concern for the teacher education component of curriculum development. The importance of teacher education appers to be clearly recognized (ASEP 1974, p. 109) but the literature in curriculum innovation and diffusion provides evidence of dissatisfaction with what is attempted in communicating ideas to teachers. Associated with this one gains a feeling of disappointment with the impact of new curricula in schools, with teacher education high on the list of possible reasons for this situation.

This concern for more appropriate teacher education in association with curriculum development is even more important when recent trends in science curricula are considered. Zoller and Watson (1974) have analyzed the characteristics of the "second generation" science projects as one method of developing guidelines for appropriate teacher education programs. This type of analysis is well worthy of examination in the Australian context if only for the reason that the six trends mentioned by the authors seem to be trends very apparent in the ASEP materials.*

As well as a consideration of the new curricula some analysis of teachers seems to be also required. Zoller and Watson (1974) provide a very brief reference to the nature of teachers as applied to teacher education for curriculum change:

> Understandably, because of his responsibility, training and experience, the teacher is not only conservative in his nature, but also brings with him a variety of attitudes and interpretations gained throughout his life. He is also under strong social pressures to remain conservative. Therefore innovaters should be prepared for the worst and not expect that brief exposure to any type of teacher training program would make major or abrupt changes. (p. 93)

If teachers are so conservative that we need to be "prepared for the worst", is it possible to develop teacher education programs that will be of any value in communicating curriculum ideas to teachers? This paper takes the view that it *is* possible and that a better understanding of the teacher (as well as the new curricula) may indicate guidelines for more effective teacher education programs.

Before turning attention to the data gathered in the study it is worth briefly considering some of the assumptions that are made by advocates of increasing the emphasis given to teacher education in curriculum development. Firstly there is an assumption that teacher education programs need to be associated with new curricula. There is a feeling that science teachers require

---

* In brief Zoller and Watson are suggesting that the new courses are tending to emphasize (I) social problems, (II) integration of the science disciplines, (III) practical themes and local problems, (IV) decision making skills for pupils, (V) interests and values (affective) objectives, (VI) less structure in the materials and more flexibility to cater for individuals and groups (pp. 94, 95).

more than teachers' guides, availability of the materials, informal discussions etc. to be able to use the materials effectively. Goodlad (1969) argues that such a process of "osmosis" is likely to be inadequate, yet Burkman (1974) attributes the "massive classroom impact" of the ISCS program to the nature of the materials.

> . . . it was decided to build the desired content and instructions rationale directly into a set of materials for the student. It was assumed that with such materials an untrained teacher could adapt to a new classroom role and to unfamiliar content while his students continued to learn more or less effectively. In other words it was assumed that well prepared student materials could serve to train teachers as well as communicate with students. (Burkman 1974, pp. 54-55)

At this stage we are involved in one of the fundamental questions for curriculum developers – the production of "teacher proof" curricula (e.g. ISCS) or the production of materials which require "curriculum proof" teachers. It is the latter approach which requires far more attention to teacher education in the curriculum development model.

A second assumption that is made when calling for a greater emphasis on teacher education is that the new curricula require changes in the classroom organization and teaching role. It may be possible for teachers to use new materials with very little change in their teaching approach. However one would hope that this follows from a careful consideration of the rationale, philosophy etc. of the new curriculum. Therefore careful consideration becomes a basic requirement for the teacher in any curriculum development program, making some emphasis on teacher education unavoidable.

A third assumption is that the changes involved in a new curriculum are worthy of consideration. Will the trends apparent in recent science curricula provide more appropriate programs for today's pupils? Again an understanding and consideration of the issues involved in this question involves us in teacher education activities far beyond what is usually attempted in Australia. This paper is therefore based on two propositions: (1) A greather emphasis on teacher education associated with curriculum development is warranted, if only to consider the important issues and trends raised by new curricula. This emphasis becomes even more urgent when this consideration leads to teachers attempting changes in their classrooms. (2) Information about Australian science teachers, at the time they are making curriculum decisions, may indicate guidelines for planning future teacher education activities.

Design of the study

The data were gathered from 151 science teachers in February/March 1974. Each of the teachers was preparing to teach Grade 7 Science in co-educational state high schools. Later information (October/November 1974) allowed identification of the teacher's use of ASEP in its first year of availability: 26 teachers were grouped as *intensive users* (i.e. they had used more than two ASEP units and used the material for more than one term (approx. 14 weeks); 33 teachers were grouped as *non-users* (i.e. no use made of ASEP materials in 1974; 19 teachers had used the materials but not to the same extent as intensive users; and 30 teachers were excluded from this study because they had volunteered to use ASEP materials for other studies. Figure 1 shows the groups of initial attributes which were examined in terms of subsequent use of the ASEP materials.

The stages in adoption set out by Havelock (1970) provided a basis for examining teachers' intentions with respect to the materials in March: 19% of teachers indicated they had no plans to use the materials in 1974, 11% planned to make up their mind after examining

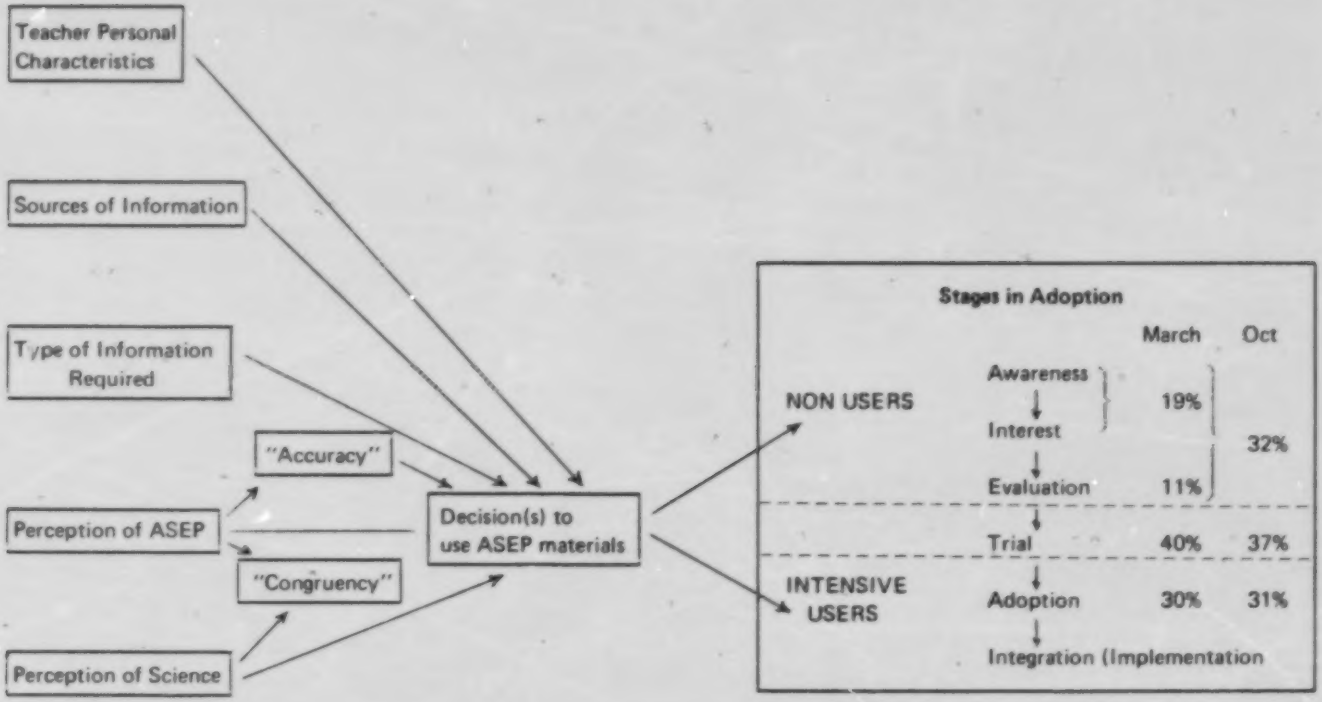


FIGURE 1: Outline of the Study

the ASEP units (*Evaluation*), 40% planned to try one unit (*Trial*), while 30% planned to use a number of units (*Adoption*). It is interesting to compare these intentions with the actual use of the materials in the first year (31% Adoption, 37% Trial). In general, teachers were able to follow through their intentions despite some initial distribution delays.

Results and possible implications

The nature of the decision. Table 1 summarises the information gathered about how the Form 1 science program was decided for 1974.

TABLE 1

Method of Making Curriculum Decisions

Nature of the Decision (Form 1 Science Course . . . 1974)

	% Teachers	Intensive Users	Non Users
Decision made by teacher concerned	19	4	9
Decision made by someone in authority	11	2	3
Decision made by science staff	40	20	13
Continuation of previous course (no decision) ..	30	-	8
	(N=151)	N=26	N=33

$\chi^2 = 8.17$ 3df.

$p < 0.05$

The evidence indicates the importance of the science staff as a group making the curriculum decisions. A survey by Beeson and Gunstone (1975) showed that 85% of science co-ordinators felt that the science staff as a group should make the curriculum decisions. Table 1 shows 40% of teachers were involved in group decisions for the 1974 programmes. For 20 of the 26 teachers who make extensive use of the ASEP materials a group decision was involved. One can suggest that courses like ASEP require the agreement and support of all science teachers in a school. There are significant costs involved and a need for teachers to co-operate in the distribution of units and apparatus between classes.

One possible implication for teacher education is to use the intact science staff of each school as the focus for in-service activities. The evidence indicates that further investigation of ways of overcoming practical difficulties, and developing activities for the science staff as a group, is warranted.

Sources of information about ASEP. From Table 2 it can be seen that teachers considered themselves to be well informed about ASEP: 79% of teachers felt they had at least enough information to consider and use the materials. Not surprisingly the eventual adopters of the materials expressed a higher level of knowledge than the non-users.

TABLE 2

Teachers' estimate of their knowledge of ASEP
(March 1974)

	% Teachers	Intensive Users	Non Users
Nil	8	0	2
Little knowledge, need to know more	13	1	8
		1	10
Sufficient knowledge to consider and use the materials	69	20	22
A great deal known	10	5	1
	(N=151)	(N=26)	(N=33)

$\chi^2 = 8.1$
$p < 0.05$

TABLE 3

Sources of Information about ASEP (February-March 1974)

Source	% Receiving Information	Teachers receiving information	
		Mean Score(Rank)[α]	% 1st Choice
Journal articles*	78	2.05(6)	38
ASEP units (Trial form)	84	2.45(3)	62
Teacher education materials	59	2.10(4)	33
Newsletters*	62	1.97(7)	27
Newspapers	44	1.31(10)	12
In-service courses	52	2.63(2)	72
Trials teacher	15	2.86(1)	87
Informal within school	70	1.85(9)	27
Informal outside school*	58	1.89(8)	26
Pre-service courses	39	2.08(5)	75

* $p < 0.05$ for t test of difference between the means of intensive users (N=26) and non-users (N=33)

α Score based on range from little use = 1 to very useful = 3.

Table 3 provides some information about how teachers obtained their information about ASEP. Teachers were asked to rate 10 sources of information from very useful to no information. The first column gives some indication of how widely distributed each source of information had been during the four years of ASEP development. It is worth noting the limited

distribution of inservice courses (52% of teachers) and teacher education materials (59% of teachers). Yet teachers seemed to be able to obtain copies of the ASEP trial units (84% of teachers) despite the limited production of these materials. Column 2 sets out a rating of each source of information after removing those teachers who received no information in each case. Experience as a trials teacher, in-service courses and examining the trial units proved to be most useful for those teachers involved in these activities, (columns 2 and 3).

In their overall ratings the extensive users initially rated journal articles, newsletters and informal contacts outside the school significantly higher than non-users. This was almost entirely due to a larger number of non-users receiving no information from journal articles and newsletters. Although more intensive users received information from contacts outside the school, they also regarded this as significantly more valuable in providing them with information.

Any implications for teacher education are probably related to the use teachers made of the ASEP trial units. Teachers seemed to have been prepared to obtain units and make up their own minds. Whether it was intended or not, the materials seem to have played a similar role in the teacher education associated with ASEP as in the teacher education associated with ISCS (Burkman 1974). Future evaluation studies will be needed before we can judge the effectiveness of the various methods of disseminating information about ASEP.

Type of information teachers required. Fourteen possible types of information about new curricula were presented to teachers to rate in terms of their importance for considering ASEP materials. The results are presented in Table 4 with the items grouped into the four areas of concern in the study.

In the overall rating of the 14 items the importance attached to knowledge about the interest response of pupils (ranked 2) is interesting. Allowing teachers to make contacts with pupils who have used the materials would seem to be worth exploring in future teacher education programs. Two practical issues, viz. "cost of materials and equipment" and "likely planning and organizational problems" are also very important types of information in teachers' minds.

Teachers who ultimately made intensive use of the materials differed from non-users in the importance attached to the "rationale and philosophy of ASEP", "theory of learning associated with ASEP" and two of the practical issues related to ASEP, (see Table 4). Tom (1973) and Gross *et al.* (1971) provide case study information which suggests that teachers develop very practical concerns when they begin using new curricula. They argue that these concerns are often very different from the type of information provided in in-service programs. The data collected in this part of the study seems to indicate that teachers are concerned with the areas of information related to Aims and Philosophy of the program. However once teachers start using the ASEP materials there is evidence that practical concerns (many of which were not anticipated) become quite important (Northfield, 1975).

If teachers' concerns do vary depending on the stage of adoption of a new curriculum, consideration may need to be given to teacher education programs continuing for a longer period of time. Prior to using the materials there may be an emphasis on aims, philosophy, initial pupil responses, etc. giving way to a more supporting function as the unique problems of using the materials become evident in each classroom.

The position prior to using the materials indicates some scope for variation in teacher education activities to suit different groups. The teachers who did not use the materials in the first year ranked general practical issues and the opinions of others more highly than the eventual users. The latter ranked aims and philosophy issues and practical issues related to ASEP more

TABLE 4

Type of Information Required by Teachers

Type of Information	Total Group	Intensive Users	Non-Users	t test difference between means
	Mean[α](Rank) (N=151)	Mean[α](Rank) (N=26)	Mean[α](Rank) (N=33)	
Aims, Philosophy of the Program				
• Rationale & philosophy of ASEP	2.31(4)	2.58(2½)	2.03(7½)	**
• Theory of learning on which the course is based	2.20(6)	2.58(2½)	2.03(7½)	**
• Aims and objectives of the materials	2.46(1)	2.73(1)	2.44(1)	
Practical Issues – General Nature				
• Costs of materials and equipment	2.30(5)	2.38(4)	2.42(2)	
• Knowledge of new subject matter	1.84 (10½)	1.68(12)	1.88(9)	
• Likely planning and organizational problems	2.32(3)	2.19(6)	2.28(3½)	
• Methods of assessing pupil achievement after using the materials	1.84(10½)	1.77(11)	2.09(5)	*
• Methods of determining the initial ability of pupils	1.63(13)	1.50(14)	1.75(11)	
Practical Issues (Related to ASEP)				
• Ways of integrating ASEP with other subjects	1.89(8)	2.08(7)	1.67(12½)	*
• Ways of organizing group work	1.71(12)	1.88(10)	1.34(14)	*
• The teaching role being suggested	1.85(9)	1.96(9)	1.78(10)	
Opinions and Information from Users				
• Test results of pupils using trial materials	1.45(14)	1.54(13)	1.59(12½)	
• Opinions of teachers who have used materials	1.98(7)	2.00(8)	2.06(6)	
• Interest response of pupils who have used the materials	2.36(2)	2.27(5)	2.28(3½)	

* $p < 0.05$ ** $p < 0.01$.

α Score based on range from no importance = 0 to extremely important = 3.

highly. Perhaps teacher education should concentrate on the concerns of the non-users. These may be the initial barriers to further involvement with the materials.

Teachers' perceptions of ASEP / perceptions of science. Herron (1970) describes a "curriculum problem of paramount importance" in the following way:

> . . . that of assessing the effectiveness of placing new curricular materials in the hands of teachers who, dedicated though they may be, do not perceive or understand the underlying philosophy, assumptions, design and purpose of those materials. (p. 49)

Herron then develops a model of curriculum innovation involving the teacher's view of the subject and the teacher's view of the new curriculum. These two views can then be related to see how well they correspond (congruency) and how well the view of the curriculum is understood (accuracy). Figure 1 shows how these variables have been incorporated in this study.

A preliminary analysis of the data used to derive these variables suggests that the Herron model may prove to be very useful in thinking about curriculum innovation and the implications for teacher education. In this paper it is not possible to develop the details of how each scale was derived but the findings are briefly stated below and shown in Figure 2.

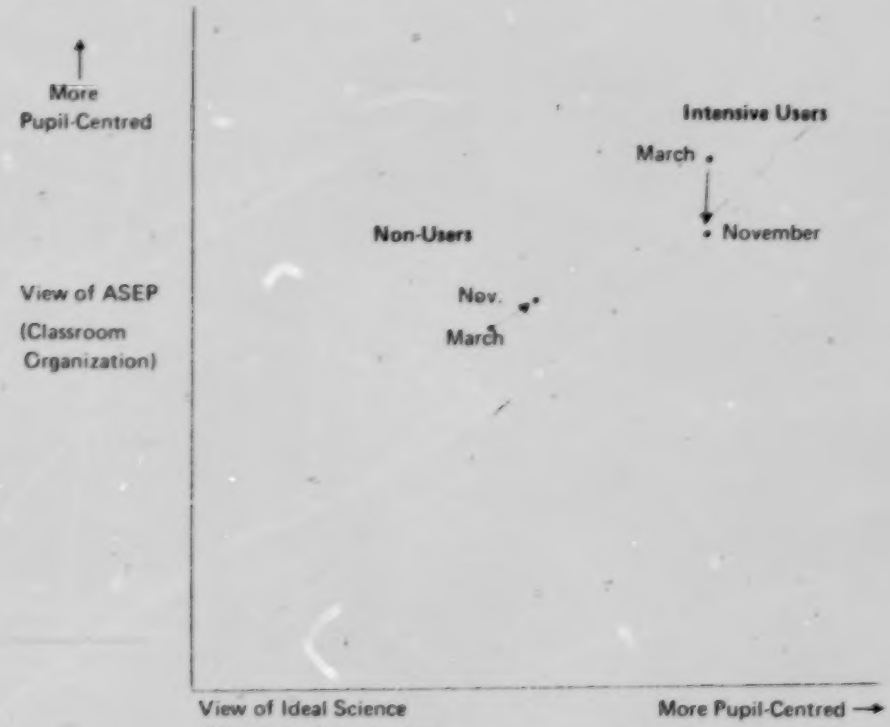


FIGURE 2: Teachers' Perceptions of ASEP and Ideal Science (Classroom organization dimension)

(i) A score on classroom organization was derived from factor analysis of the responses to a questionnaire. Prior to the availability of ASEP materials the teachers who made intensive use of the materials were significantly more pupil centred in their view of science and had a more pupil centred view of ASEP, (t test significant at 0.01 level).

(ii) Both intensive users and non-users regarded the ASEP course as more pupil centred than their ideal science course, (see Fig. 2), but it is interesting to note the shift in the intensive users' view of ASEP after one year. Their view of the ASEP materials shifts to a less pupil centred view while their view of science has remained almost the same.

In evaluating the impact of ASEP program on teachers one may be a little disappointed to find little overall change in the intensive users' perception of their ideal science course after using the materials.

(iii) A measure of accuracy of teachers' perceptions of ASEP was obtained by comparing the teachers' responses to items with the agreed response of "experts" familiar with the ASEP materials. Prior to the availability of the materials the intensive users had a significantly more accurate view of the ASEP curriculum than the non-users, (t test significant at 0.01 level).

(iv) A comparison was made between each teacher's responses to the same items for ideal science and for ASEP materials. This provided an estimate of the congruency between view of science and view of ASEP. Intensive users had a far more congruent view (particularly on items related to classroom organization) than teachers who subsequently did not use the materials, (t test significant at 0.001 level).

The last two results are certainly not surprising. Most of what is done in teacher education programs is designed to give teachers an accurate view of the new curriculum. However starting with a person's view of the subject and attempting to relate the new curriculum to this already formed perception is not often attempted in in-service programs. With flexible programs such as ASEP it becomes feasible to talk in terms of the way the materials may be used to suit particular teachers in their situations, i.e., attempt to achieve congruence with individual teachers' view of science.

Summary

The nature of recent science curricula requires a careful examination of the teacher education required. The possibility of teachers following set programs to gain an understanding of a new curriculum is becoming more unlikely. In the case of ASEP it is difficult to describe one approach, sequence of units, or method of classroom organization. The selection of units and options within units requires teacher skills which go beyond an understanding of one curriculum. Our teacher education programs face new problems. Possible guidelines may be found from an analysis of the recent science curricula and a better understanding of how our teachers are operating when these new curricula become available. This preliminary analysis of teachers' responses to ASEP in its first year suggests the following as worthy of attention in teacher education programs.

1. Attempting to develop activities to involve the intact group of science staff of each school where possible.

2. Provision of copies of units for teachers to form their own opinions about possible future use.

3. Attempt to provide more opportunities for informal contact outside the school. This may be possible as teacher centres become established. It may also be possible to arrange visits

to other schools so that teachers can talk with pupils who have used the materials (a source of information rated very highly by teachers).

4. Evidence suggests that different groups of teachers have different needs at varying stages in the adoption process. Teacher education should be a continuous activity, dealing with the aims, philosophy and underlying theories of the curriculum in the initial stages and moving into a more supportive function as the curriculum is used.

5. Attempting to achieve congruence between a teacher's view of science and the view of the new curriculum. This involves starting with the teacher and exploring the ways in which the curriculum can best be used. This can be done only when the teacher clearly understands the principles on which the curriculum has been developed. There must be some opportunity for the basic ideas of a curriculum to challenge (and perhaps alter) a teacher's view of science.

The ultimate test of the impact of a new curriculum (and the teacher education program associated with it) is seen in the classroom. Therefore the way in which the curriculum is implemented (see Figure 1 – final stage in adoption) will provide the most interesting information about how ASEP has influenced our science classrooms.

References

AUSTRALIAN SCIENCE EDUCATION PROJECT. *A Guide to ASEP.* Victorian Government Printer, 1974.

BEESON, G.W. & GUNSTONE, R.F. "The teacher's role in curriculum decisions." *Australian Science Teachers' Journal*, **21** (*1*): 5-19, 1975.

BURKMAN, E. "An approach to instructional design for massive classroom impact." *Journal of Research in Science Teaching*, **11** (*1*): 53-59, 1974.

GOODLAD, J.I. "Educational change: a strategy for study and action." *The National Elementary Principal*, **48** (*3*): 6-13, 1969.

GROSS, N., GIACQUINTA, J.B. & BERNSTEIN, M. *Implementing Organizational Innovations: A Sociological Analysis of Planned Educational Change.* Basic Books, New York, 1971.

HAVELOCK, R.G. *A Guide to Innovation in Education.* Institute for Social Research, The University of Michigan, Ann Arbor, Michigan, 1970.

HERRON, M. "On teacher perception and curricular innovation." *Curriculum Theory Network*, **7** (*1*): 47-52, 1971.

NORTHFIELD, J.R. "Some general problems in using ASEP materials." *Lab Talk*, **19** (*1*): 12-13, 1975.

RAMSEY, G.A. "Curriculum development in secondary school science, Part 2." *Quarterly Review of Australian Education*, **5** (*2*), 1972.

TOM, A. "Teacher reaction to a systematic approach to curriculum implementation." *Curriculum Theory Network*, **11**: 86-93, 1973.

ZOLLER, V. & WATSON, F.G. "Teacher training for the 'second generation' of science curricula: the curriculum proof teacher." *Science Education*, **58**: 93-103, 1974.

A REVIEW OF CURRENT AND PLANNED RESEARCH CONCERNING THE AUSTRALIAN SCIENCE EDUCATION PROJECT

John M. Owen

Introduction

The Australian Science Education Project (ASEP) has provided extensive curriculum materials for use in secondary schools. The sale, by the Victorian Education Department's Stores Branch, of materials valued at over $500,000 during the first year of marketing gives some indication of the extent of the impact these materials are having on science teaching in Australia. The production and widespread use of the units has provided a unique opportunity for Australian research workers interested in the problems of science education to carry out investigations concerned with the teaching and learning of science through the use of these new curriculum materials.

This paper summarizes research work which is under way or is planned into the use of the ASEP materials and procedures. In addition, the information collected during the production of ASEP, and now stored by the Curriculum Development Centre in the ASEP files, is outlined and summarized.

Furthermore, this paper seeks to inform investigators of other data which might be useful in their current studies or in new projects which they are planning. It also makes some recommendations concerning areas in which further research is urgently needed.

Research studies

In preparing this review of research in the field of science education in Australia it was decided to concentrate on the collection of information about studies specifically involving the use of ASEP materials, rather than to include more theoretical papers concerned with the problems of evaluation. Nevertheless, there have been some important contributions on evaluation in the field of science education by Australian writers, which are too valuable to be overlooked. Papers by Ramsay (1971), Baumgart (1972) and Cohen (1973), fall into this category.

Information for this paper regarding studies of the Australian Science Education Project and its materials has been gathered by:

(i) an examination of journal articles,
(ii) sending letters to researchers asking for information of their work, and
(iii) visiting colleges and universities where research in this area is under way.

An examination of the studies involving ASEP and its materials lead to classifying them under the following categories: production of materials; the use of materials in schools; teacher reaction to ASEP; student reaction to ASEP; and "other projects". Within these major categories there are clearly more specific areas of research. A summary of research is presented in the appendix.

Production of materials. During the production of ASEP materials, extensive formative evaluation studies were undertaken by members of the ASEP team. Trials were carried out in Victoria and on a national basis. The information collected led to the modifications of the units before publication. Further details of the data available from this evaluation are presented in a later section

of this paper, and it is to be regretted that no account of this work has been published.

The production of the ASEP units provided an opportunity for an independent team to analyse the development and evaluation procedures used in a large curriculum project. Cohen and his co-workers visited the ASEP offices in 1971 and 1972 and investigated the procedures being used during the formative evaluation stage. This study involved interviews with members of the production team and the completion of questionnaires by them as well as an examination and analysis of data in the ASEP files. As yet a report by these workers of their findings has not been completed, due to dispersion of the research team, although a draft report was prepared and presented to the ASEP staff.

Use of materials in schools. Studies which report the use of ASEP materials have, in general, been confined to surveys conducted within one state. The only Australia-wide survey has been carried out under the direction of Mr L.G. Dale by the Melbourne Office of the Curriculum Development Centre. It has been concerned with the use of curriculum packages and textbooks in all areas of science in Australian secondary schools. The schools were sampled in a systematic way and replies received from 99 of the 130 schools approached in late 1974. No attempt was made to determine the extent of use of the various materials. The report showed that at grades 9 and 10 ASEP was used in a greater percentage of schools than any other curriculum materials, but that at grades 7 and 8, the Junior Secondary Science Project (JSSP) materials were the most popular. This report has not as yet been published.

Page, a science consultant for the Education Department of South Australia, sought information on the use of ASEP materials in schools in that state. In an initial report (Page, 1975) showed that of the replies received from 63 schools, 54 percent had already used ASEP, and that all except one were planning use in 1975. Information was also reported on the relative use of units, and on the problems and needs of teachers when ASEP was used in schools. A further report is planned. In addition, the South Australian Science Teachers Association is conducting a survey into the availability of ASEP materials, and their effect on the planning of school science syllabi in South Australian schools in 1975 (South Australian Science Teachers' Association, 1974).

In a survey which attempted a detailed description of the activities of science teachers in Victorian schools Gunstone and Beeson (1974) obtained information on the use of curriculum materials. In March 1974, questionnaires were sent to 214 schools, and 114 replies were received. The report showed similar results to the Curriculum Development Centre survey regarding use of JSSP and ASEP materials in Forms I-IV and also indicated that "teacher developed materials" were often used at all levels from Form I to Form IV. Again, *amount* of use was not reported.

Northfield, (1975a, 1975b) as part of a larger study, sought information concerning teacher knowledge of ASEP by gathering data from 150 Form I teachers in Victorian secondary schools. As well as obtaining evidence about ways in which a knowledge of ASEP materials is acquired, information was gathered about the reactions to some of the problems teachers encountered in using the ASEP units.

Fisher and Fraser (1974) conducted a survey of teacher opinions amongst a group of Tasmanian teachers after using ASEP materials in 1973. The opinions of 38 teachers in 38 schools have been examined. Some of the more clear cut results indicated that the use of ASEP materials did not reduce job satisfaction and that units were not as suitable for slow learners as for students with higher intelligence. The most favoured reasons for use of ASEP were "pupil interest" and "relevance to the school environment". The average figure for the percentage of future science class time to be devoted to ASEP use was 51% at Form I, grading to 31% in Form IV. Burr is also believed to have completed a study to discover what science curriculum materials

are being used in Queensland schools, but no details were available.

Teacher effects. Research on teacher values and behaviour has been carried out using ASEP units. One continuing study was outlined by Power (1972). The research involved not only former ASEP staff but also practising teachers and Diploma in Education students at the University of Queensland. Before the students had any interaction with ASEP materials various clusters of teachers and students were identified on the basis of their values about teaching, and the agreement of values they held with those expressed in the ASEP literature (Power and Tisher, 1973(a)). A further investigation examined the relation between the activities of teachers in classroom and their attitudes (Power and Tisher, 1973(b)). Selected teachers and student teachers were subjected to various training procedures which varied according to the extent of their commitment to the ASEP philosophy. The resulting analyses showed that for Diploma in Education students, both training methods and teacher values affected attitudes and achievement of students in their classes. The values held by experienced teachers, but not the training they undertook, affected pupils attitudes to science lessons. However, when the achievement of the students was examined, the reverse result was obtained (Power and Tisher, 1973(a)). A replication study using a second unit of ASEP materials has been carried out by these workers, and a report is currently being prepared.

As well as the collection of data about teacher knowledge of ASEP materials, Northfield at Monash University has carried out a study involving the use of ASEP. Teachers were asked to teach an ASEP unit and their reactions to the use of the unit were collected after the teaching was completed. As a further component of the study, student attitudes to ASEP were measured using the Science Classroom Inventory administered before and after the experimental procedure. A detailed report of findings is in preparation.

Student effects. Fraser (1973) at Monash University has been preparing content-free tests to measure inquiry skills. He has also used a Test of Understanding Science and attitude questionnaires. This investigation has compared classes using only ASEP materials and control groups in Melbourne high schools in 1974, and a full report is being prepared by Fraser. In addition, the tests used in Fraser's study were applied to students in eight Form I classes in Tasmanian high schools. Change in performance on all criteria was studied as a function of socio-economic status of student (measured by father's occupation), IQ and sex. Results showed that the major difference in performance occurred between students of high and low IQ, in favour of more intelligent students (Fraser and Fisher, 1974).

Trebilco (1974) has examined the development of activity skills arising from the use of the ASEP unit *Made to Measure.* A written and a practical test were used, and pre- and post-tests were administered to 87 students in 4 Victorian trials classes. Slight increases in scores from the pre-test to the post-test were observed for seven of the thirteen objectives. However, these increases in score were not statistically significant and it was not surprising that some of the results of the practical test were contrary to results on pencil and paper practical items.

Little of the material gathered during the formative evaluation phase, has been published, due mainly to the limitations of time and resources available to staff. However, Pascoe and Shepherd (1974) have presented a short report of student ratings on thirty-one units. When trialling occurred, opinions of students were sought on various aspects of ASEP materials. The results indicated that if students enjoyed the units as a whole the various aspects were also rated highly. The effects of a global judgment serve as a warning for those who seek to obtain information on the short-comings of a unit from an examination of student opinions.

Tuffin, a Master of Education candidate at the University of Western Australia, has examined the attitude changes of students towards science for students subjected to two terms of ASEP teaching. The test used was one developed by Hedley at Michigan State University. Other data collected included student and teacher opinions of individual units. A thesis on this work is about to be submitted.

McKenzie in Queensland has received a grant from the Australian Schools Commissions' Innovations Program, and is currently planning a series of evaluation studies.

Jarrett, at Flinders University, is planning a study on interactions between students in science classrooms. A comparison is to be made between students in ASEP and more formal classrooms.

Moritz at La Trobe University is in the process of preparing an instrument to compare the class environments of student groups who are using ASEP materials with those in science classes not using ASEP. It is planned to develop a Learning Environment Inventory for use in Form I classes in Victorian schools.

Other projects. There has been some research into the structure of ASEP units. Investigations of this nature have sought to identify the underlying pattern of relationships which were seen to be essential for the unit to be successfully used by teachers and students. The core of the ASEP unit *Forces* was studied in its trial form by Carss and Clarke (1972) using content analysis techniques. After identifying sequences requiring "reconstruction", the original and modified versions were trialled in schools. Clarke (1973) reported that students performed better on tests of achievement after studying the rewritten units. The study also reported that teachers still had an effect on the performance of students in the classes tested. More recently Carss has been further examining content analysis methodology as a pre-evaluative tool in curriculum development using some ASEP materials.

Kings (1974) has reported work in progress on a project which involves the translation of an ASEP unit into other languages. It is intended to have materials available in English and in several foreign languages in order to facilitate the learning of science by migrant students and to improve their level of reading comprehension in their native language. Evaluation of the use of the materials developed is continuing.

Information available in Curriculum Development Centre files

Documents relating to the production of ASEP materials have been stored at the Curriculum Development Centre in Melbourne. Two officers are employed on matters concerning science teaching, one of whom was a member of the ASEP team. This section describes the materials held by the CDC.

Guidelines Conference documents. The Guidelines Conference (18-23 January 1970) was a forum at which educators and scientists presented views concerning the proposed project and its effects on science education in Australia.

Position documents describe the manner in which the formal organization and co-ordination of the project proceeded. About eighty documents are included in the file.

Developers copies of trials materials are master copies from which modifications to units were effected as a result of all information gathered in trials. About 20 developers copies are stored at the CDC offices.

Unit folios include specifications from which trials materials were drawn up, the resulting trials copies, and a summary of evaluation statistics. The summary lists the schools used, names of outside consultants, and information received from State Advisory Committees and ASEP personnel involved in school visits. Also, ratings on aspects of the unit by teachers, consultants and observers are reported. The ratings are concerned with the prediction of the success of the unit with respect to its general structure, materials and philosophy. In most cases, the results of pre- and post-tests have been included, and comparisons between control and experimental groups have also been recorded. The student tests are often also included.

Copies of all ASEP publications used in the formative evaluation of ASEP. These comprise:

(a) a summary of position documents,
(b) teacher information questionnaires (1a, 2a, 3a),
(c) unit evaluation questionnaires (1a, 2a, 3a, 3b, 4a, 5a, 5b), which contain information about student activities, the reactions in the home and school, and student reactions to various of the ASEP units.

Areas of further research

Moritz (1972) set out some suggested research areas associated with the ASEP materials. Although he was concerned with evaluation of a formative nature, many of the topics are relevant to research after the completion of the production of the ASEP materials. Research is under way in some areas suggested by Moritz, e.g. in the area of testing students' understanding of science and describing the administration and organization of ASEP, but there are many other areas where work could be done. As an example, there appears to be some concern amongst teachers that students find difficulties carrying out science activities which require them to read and understand the contents of the student texts. When ASEP units were produced, care was taken to set the reading level so that most students at a given form level could cope with the texts. Some recent studies and informal feedback suggests there are still problems concerned with these activities. The problems of comprehension and reading levels cuts across subject boundaries, and with more emphasis being placed on student centred activities in schools, an investigation of these problems is of importance.

Cohen's work will be of interest when completed, and depending on its scope, there may be a place for a further investigation of the project in the period between May 1972 and its termination. The position documents would be useful in providing data for such a study. Moreover, the availability of these documents and guidelines conference notes should prove useful if a study were undertaken of the factors contributing to the creation and demise of the Australian Science Education Project.

An area of basic concern in any curriculum project is to determine whether the aims of the project have been achieved. Some ideas about the objectives of the ASEP project can be obtained from reading *A Guide to ASEP* and other publications relating to the preparation of materials. The ASEP materials place an emphasis on the development of inquiry skills. This raises the questions as to whether through the use of these materials over a long term students develop the ability to work effectively and independently on investigations that are open-ended.

The development of the ASEP units was based in part on an environmental approach. Linke has suggested (in correspondence) that a content analysis could be undertaken to discover thematic lines underlying units, or at a more general level to develop materials to test whether environmental issues are better understood as a result of learning with ASEP materials.

Some investigations might focus on the ways in which intending science teachers are trained in Australian universities and colleges. What are the relative advantages and disadvantages of the various forms of teacher training when the changing role of the science teacher is considered? Some research suggests that student-teachers must be introduced to various methods of handling science lessons before they enter schools if a degree of flexibility is to be retained during their careers.

The place of science in the curriculum of secondary schools now appears to be under question, despite the improved qualifications of science teachers and the production of materials designed to make science teaching and learning more exciting and enjoyable. Although more students are staying on for a longer time at school, it has been shown that smaller proportions are progressing to study science at upper secondary school levels (Lee Dow, 1971). The availability of ASEP materials provides an opportunity to investigate the approaches schools make to the introduction of a new set of materials, to problems which arise in the introduction of a curriculum innovation and the effects these materials have on the attitudes of students to science. Any study of this nature must take into account the relative influence of State Government Education Department policies, school policies, the science department as an organization and the ways in which individual teachers fit into this organization. Such a study would be concerned with the mechanisms by which curriculum planning is initiated, the factors which determine the extent of the use of curricula and why they succeed or fail. The results of such a study could have important implications, not only for the future modification and extension of ASEP materials, but also to answer more general questions about the place of organized curricula, the demands of such curricula, and the structures needed in science departments to cope with general problems of science teaching.

APPENDIX

A Summary of Current and Planned Research Concerning ASEP

Area of Research	D.C.[a]	C.[b] or P.	Researcher(s)	Place of Contact	Summary of Research
Production of materials	72	P	Cohen	Macquarie University	Examination of processes
Surveys on use of ASEP	75	P	Burr	Mt Gravatt Teachers College	Use of curriculum materials in Queensland science classes
	74	C	Dale and Greenhall	C.D.C. (Melbourne)	Science materials in use in Australian schools
	73	C	Fisher and Fraser	Tasmanian College of Advanced Education Monash University	Teacher opinions after
	74	C	Gunstone and Beeson	Monash University	Curriculum choice in Victoria
	74	C	Northfield	Monash University	Preparedness of science teachers for ASEP
	74	C	Northfield	Monash University	General problems of using ASEP
	74	P	Page	Science Consultant S.A. Education Dept.	Use of ASEP problems and needs of S.A. teachers
	74	C	S.A.S.T.A.	South Australia	Problems of ASEP supply
Research concerning teachers	73	P	Northfield	Monash University	Changes of teacher opinion about their science classes after using ASEP
	72	P	Tisher and Power	Monash University and University of Queensland	Effects of teacher training on teacher attitudes to science teaching

Area of Research	D.C.[a]	C.[b] or P.	Researcher(s)	Place of Contact	Summary of Research
Research concerning students	72	P	Fraser and Fisher	Monash University and Tasmanian College of Advanced Education	Test of enquiry skills Test of understanding science Test of attitude changes
	75	P	Jarrett	Flinders University	Interaction effects in ASEP and 'formal' classrooms
	74	P	McKenzie	Wavell State High School Queensland	Student characteristics including use of motivation measures
	75	P	Moritz	La Trobe University	Learning environment
	70	C	Pascoe and Shepherd	Curriculum and Research, Victoria	Ratings of aspects of ASEP units
	74	C	Trebilco	Rusden State College, Victoria	Activity skills
	73	P	Tuffin	University of Western Australia	Attitude changes to science over 2 years
Other projects	72	C	Clarke and Carss	STEP Research Officer and University of Queensland	Logical structuring of an ASEP core
	74	P	Kings	Rusden State College, Victoria	Translation of ASEP units into foreign languages and evaluation of use in schools

a. Date commenced.

b. Status: Completed (C) or in progress (P).

References

BAUMGART, N. "A model for curriculum evaluation." *Research 1972*, 1-12, 1972.

CARSS, B.W. & CLARKE, J.A. "Content analysis of an ASEP unit: 'Pushes and Pulls' examined." *Australian Science Teachers Journal*, **18**(*4*): 59-61, 1972.

CLARKE, J.A. "Ausubel and ASEP: an application of cognitive field learning theory to an ASEP unit." *Australian Science Teachers Journal*, **19**(*4*): 92-97, 1973.

COHEN, D. "Evaluation reinterpreted." *Australian Science Teachers Journal*, **19**(*2*): 29-34, 1973.

FISHER, D.W. & FRASER, B.J. "A survey of teacher opinion after using ASEP in Tasmania." *Australian Science Teachers Journal*, **20**(*1*): 89-96, 1974.

FRASER, B.J. "A test of inquiry skills." *Science Education: Research 1973*, 77-94, 1973.

FRASER, B.J. & FISHER, D.W. "Changes in pupils during a year of ASEP in Tasmania." *Australian Science Teachers Journal*, **20**(*2*): 90-100, 1974.

GUNSTONE, R. & BEESON, G. "Curriculum choice in Victoria." *Lab Talk*, **18**(*4*): 15-27, 1974.

KINGS, C.B. "Evaluating bilingual teaching materials." *Research in Science Education*, **4**: 89-96, 1974.

LEE DOW, K.C. *Teaching Science in Australian Schools*. Melbourne University Press, Carlton, 1971.

MORITZ, K.W. "Some suggestions for research related to ASEP." *Research 1972*, 18-27, 1972.

NORTHFIELD, J. "Some general problems in using ASEP materials." *Lab Talk*, **19**(*1*): 12-13, 1975(a).

NORTHFIELD, J. "ASEP . . . Are teachers prepared?" *Australian Science Teachers Journal*, **21**(*1*): 21-25, 1975.

PAGE, R. Junior Science Newsletter, Education Department of South Australia, 1975.

PASCOE, R. & SHEPHERD, R. "Do student opinions really matter?" *Australian Science Teachers Journal*, **20**(*1*): 97-99, 1974.

POWER, C.N. "The Australian Science Education Project as a stimulus for research: a progress report of a study of teaching strategies used with ASEP materials." *Research 1972*, 13-20, 1972.

POWER, C.N. & TISHER, R.P. "Teacher values and their association with teaching strategies in ASEP classes." *Science Education: Research 1973*, 51-60, 1973(a).

POWER, C.N. & TISHER, R.P. "Congruency with ASEP: a study of practices and values of a group of teachers and Dip.Eds." *Australian Science Teachers Journal*, **19**(*3*): 84-88, 1973(b).

DETECTIVES ARE BORN, NOT MADE
or
TWO SIDES OF A CHEMISTRY LEARNING COIN

Peter J. Fensham & Kaye Nickless

One of the classic activities of chemists has been the identification of "unknown" substances. In general these unknowns fall into two categories. The first group consists of those substances, the existence of which is evident from their properties, but for which the structure has not yet been determined. The second group are substances the details of which are known, but are now being identified afresh in mixtures or in other chemical problem situations. Spectacular examples of the identifying process of chemical "firsts" are the work of Woodward *et al.* on chlorophyll and of Watson, Crick and Wilkins on DNA. The uniqueness of each of these identifications means that they are more associated with the education of postgraduates in research methods.

The skills involved in the determination of the second group can be considered in two directions. These are essentially represented by the equations *I* and *II*:

$$\text{Known Substance of Known Structure} \xrightarrow{E} \text{Data} \quad \ldots\ldots\ldots\

rather like a detective who has some clues which relate to a person who has a real but initially unknown identidy. In *I* he is more like a portrait painter, biographer or photographer who knows a person and is seeking to provide for others a representation of him. Detectives are not automatically good biographers, or vice versa.

Relative difficulty

A set of tests was devised on the interpretation and establishment of mass spectra for organic compounds. The compounds and their spectra were chosen in the following way. Four, A, B, C and D, belonged to the same class of compounds and gave a number of peaks that differed only by ±1. The other two, E_I and E_{II}, were heterocyclics with an equal number of peaks.

Fifteen final year chemistry students undertook the test after completing a lecture course on organic mass spectrometry. Each did six questions – 2 of A, B, C, D for *I* and the other 2 for *II*, and E_I for *I* and E_{II} for *II*. The time taken for each question was listed and at the end the students were asked to draw up an order of difficulty.

The results are given in Table 1.

TABLE 1
The number of "correct" responses and the average time for test items of types *I* and *II*

Cpd.	*I*		*II*	
	Aver. Time min.	No. "Correct"	Aver. Time min.	No. "Correct"
A	4.6	$^6/_8$	9.6	$^6/_8$
B	6.0	$^5/_7$	11.8	$^4/_8$
C	5.3	$^6/_8$	18.3	$^4/_7$
D	3.7	$^6/_7$	13.8	$^0/_8$
E_I	6.5	$^{10}/_{15}$		
E_{II}			13.5	$^0/_{15}$

A response was rated as correct if the majority of the peaks were given for items of type *I*, and if the compound was identified in the type *II* items.

All students reported that they found the items of type *II* generally more difficult than the items of type *I*.

Although the number of students is small and the testing suggested a number of procedural weaknesses, the direction of difficulty seems quite clear. The average times, the correctness of the responses and the students' own perceptions all indicate that the problem solving in *II* is considerably more taxing than that in *I*. At least in this area, chemical detecting is more difficult than chemical portrait painting.

The skills of *//* or chemical detecting

The spectra of a compound, methyl cinnamate, was chosen for the investigation of the human skills, *//*. It had been suggested by Professor Morrison of La Trobe University, because at one stage he had more success in teaching a computer to interpret this spectra than he had with his students.

Six research students were invited to interpret the spectra and to provide as full a vocal account as they could of the processes involved. These six recorded solutions were then transcribed and compared. From them it was possible to set out a sequence of processes which encompasses, allowing for iterations and some variations, the methods of the various solvers. This sequence is shown in Figure 1.

The sequence in Figure 1 does not indicate how the processes – determine, ascertain, check and rule out, etc. – were in fact carried out. Indeed, answering the deceptively simple question, what size peaks can I neglect?, was an important and difficult step in the solution. Certain boundary conditions emerged for the schematic solution of Figure 1. Stepwise fragmentation only is used, the appearance of a molecular ion is assumed and the resolution of *cis* and *trans* isomers is not expected.

One research worker whose solution closely paralleled the sequence was then given one of the tests. All items were correctly answered and his times for *I*: B, D, E and *//*: C, E, A were 7, 3, 10, 50, 15 and 3 minutes respectively.

Teaching the skills of *//* or chemical detecting

The sequence of processes was used to develop flow charts from which frames of a learning program for the determination of a compound from its spectra could be written. Several flow charts consistent with the sequence were possible but the one chosen involved loops which enable the steps to be repeated and additional ones added to re-inforce the main idea if mistakes are made. The program can be regarded as made up of five sequential chapters: 1. Determination of molecular ion, 2. Determination of possible fragments in stepwise order, 3. Construction of possible molecular structures, 4. Fragmentation and checking of structures, and 5. Postulate best structure.

Chapter 3 turns out to be of particular interest. It is at this stage of their problem solving that most of the student failure occurred in the tests of difficulty described earlier. It is also a difficult section of the learning program for which to write frames.

If the program is written with a specific example in mind, it is comparatively simple to hint in the frames at a range of possible structures that include or even highlight the ultimate structure. However, if the program is to be written in a form that works from a quite general or "unknown" set of molecular fragments, then the writing of these frames, both questions and answers, becomes a problem of great difficulty. Nor is much help to be gained from those computer programs that exist for the process, *//*. The sequence in Figure 1 is very like the flow chart of the computer program developed by Crawford and Morrison (1971), except that the latter incorporated several sub-programs. One of these contained a list of key peaks and their characteristics. The operation in this sub-program consisted of rapidly checking through a list. Another example of such an "inhuman" process is to be found in the computer program of Lederberg *et al.* (1969) and Duffield *et al.* (1969). They have provided a general approach to the problem *//* by making use of DENDRAL, a program which generates organic compounds. The first step in their method is to make preliminary inferences by comparing key peaks from the spectra with

FIGURE 1: The sequence of processes in determining a compound, methyl cinnamte, from its spectra.

the data bank provided by McLafferty. This step of preliminary inference provides the functional groups for which possible structures are then generated by DENDRAL. Each of these is then fragmented and checked against the observed spectra, to refine the possible structures.

Crawford and Morrison (1968, 1969) describe several methods for beginning the identification. Each one compares features of the experimental spectra with data banks which contain condensed spectra, or specific peak probabilities or hypersphere cluster centres which are derived from spectral data for a number of compounds of the same class.

In all these examples in which computers are used for chemical inference, there is built into the program, in one way or another, knowledge (and lots of it) of the solutions of problems of type *II*. In other words the program for a type *II* solution requires type *II* solutions en route.

Careful analysis of the activities of the human solvers of problem *II* suggested the same dilemma. For example, when they constructed possible molecules, comments occurred like, "When I had a peak (or combination of peaks) like that once before it turned out to be an X, so I'll try an X again", or "This reminds me of a spectra I once solved for Y so it's probably rather similar with some branching".

Discussion

The comparison of human failure, human success and computer "success" at the skills of *II* appears to have led to the finding that only by knowing how to do them can one learn them. This depressing conclusion for those without the skills is, of course, nonsense since there are many individuals who once did not have the skills and have yet acquired them.

How then can this happen? The answer must lie in the risks and short cuts human problem solvers take. Since it is not possible for a human solver to use the very laborious library searches, or "all possibilities" approaches that are used in the computer solutions, he has to put forward a "best estimate" and hope it won't lead into a morass of ambiguity and contradiction. Another comment from the human solvers was, "it's most likely to be something like W, so I'll try W^{I} and then W^{II} if W^{I} doesn't work". The basis for these estimates, or acts of intuition, may very well be general wisdom (specific knowledge plus appreciation of the likelihood of occurrence in chemistry). This will lead to lower risk and fewer mistakes from which recovery is not possible. It will also be a guide concerning when a proposed structure should be abandoned.

Licklider (1965) has suggested in discussing the man-computer partnership that man excels in heuristic phases of problem solving while the computer excels in executing explicit algorithmic procedures. This distinction is useful although it can become rather blurred in problems with the complexity of the present type.

The skills in *II* involve a series of estimates and evaluations – both high level cognitive activities. Those in process *I* are in general very much simpler – the application of an explicit procedure. The two tasks can be well represented by the branching paths shown in Figure 2.

In going from A to B (process *I* at each junction the figure suggests that there is an obvious route of the two possibilities. Only a "perverse or stupid" person would get lost. However from B to A each junction presents a real choice and a mistake must be avoided or checked before yet another choice appears.

The action of the computers at each junction is very laborious (although quite rapidly done) and has the problem that they make each junction or choice much more complex than it really is.

Perhaps the best procedure to teach the skill *II* is to develop a series of graded problems

FIGURE 2: A representation of the two problems *I* and *II*.

so that the right branch in Figure 2 stands out so obviously that the learner finds he has acquired the skill almost without facing the problem in all its complexity. This limited achievement thus, for a more difficult example, becomes for him each time a guiding counterpart of the built-in lengthy solutions that the fast but inefficient computers use.

In this way we may enable more students to get around the barrier for which this study provides evidence – Chemical detectives are born, not made.

References

CRAWFORD, L.R. & MORRISON, J.D. "Computer methods in analytic mass spectrometry: identification of an unknown compound in a catalog." *Analytical Chemicsty*, **40**: 1464-1469, 1968.

CRAWFORD, L.R. & MORRISON, J.D. "Computer methods in analytic mass spectrometry: empirical identification of molecular class." *Analytical Chemistry*, **40**: 1469-1474, 1968.

CRAWFORD, L.R. & MORRISON, J.D. "Computer methods in analytic mass spectrometry: development of programs for analysis of low resolution mass spectra." *Analytical Chemistry*, **43**: 1790-1795, 1971.

DUFFIELD, A.M. *et al.* "Applications of artificial intelligence for chemical inference II. Interpretation of low resolution mass spectra of ketones." *Journal of the American Chemical Society*, **91**: 2977-2981, 1969.

LEDERBERG, J. *et al.* "Applications of artificial intelligence for chemical inference I. The number of possible organic compounds. Acyclic structures containing C, H, O and N." *Journal of the American Chemical Society*, **91**: 2973-2976, 1969.

LICKLIDER, J.C.R. "Man-computer partnership." *International Science and Technology*, May, 18-26, 1965.

COGNITIVE PREFERENCES AND ACHIEVEMENT IN PHYSICS, CHEMISTRY, SCIENCE AND MATHEMATICS

Lindsay D. Mackay

Cognitive preferences and their relationship to Ausubel's model of meaningful verbal learning

Ausubel's model for the learning and retention of meaningful verbal material assumes the existence in the learner of a heirarchically organised cognitive structure in which the organisational principle is that of progressive differentiation in degree of generality and inclusiveness. Cognitive structure can be defined as "an individual's organisation, stability, and clarity of knowledge in a conceptual area or subject-matter field relative to meaningful new learning tasks in this field", (Ripple 1964, p. 190). This cognitive structure is assumed to be organised in terms of highly inclusive conceptual traces under which are subsumed less inclusive subconcepts and specific items of information. New learning material is potentially meaningful if it is relatable in a non-arbitrary substantive manner to relevant concepts in the particular cognitive structure of a particular learner, *i.e.*, the cognitive structure includes the requisite intellectual capacities, ideational content, and experiential background. Meaningful learning occurs when potentially meaningful material enters the learner's cognitive field, interacts with his existing cognitive structure, and is incorporated into the cognitive structure by subsumption under more inclusive conceptual traces in accordance with the principle of progressive differentiation. Actual meaning emerges from the interaction of potentially meaningful material with stable elements in the learner's cognitive structure which enables the establishment of significant meaningful relationships with pre-existing elements in the learner's cognitive structure.

In this view, details of a given discipline are learned as rapidly as they can be fitted into a conceptual framework consisting of a stable and appropriate body of general concepts and principles.

In this model of learning, because of the importance attributed to the interaction of new learning material with the cognitive structure of the learner, the learning and retention of meaningful verbal material are pictured as greatly influenced by the attributes of the particular concepts available in the cognitive structure of the learner, and the way in which the cognitive structure is organised. In fact, existing cognitive structure is regarded as the major factor affecting meaningful learning and retention.

Specifically a learner's cognitive structure is hypothesised to be involved in the meaningful learning and retention of new materials in a number of ways:

- *potential meaningfulness* depends upon the availability of relevant subsuming concepts in the learner's cognitive structure;
- *conversion of potential meaning into actual meaning* occurs through the interaction of new learning material with relevant elements of the learner's cognitive structure;
- a learner's cognitive structure is *the repository of the outcomes* of all previous experiences;
- the learner's cognitive structure is the *site at which obliterative subsumption occurs*, and as a result of which materials lose their dissociability from other elements, and forgetting occurs;
- *a learner's potential to meaningfully learn* new material depends upon the organisational properties of the learner's knowledge in the relevant subject area *i.e.*, upon the

learner's cognitive structure; and

- *learning strategies* form part of a learner's cognitive structure (Ausubel, 1963).

In terms of Ausubel's model, a learner's cognitive structure should be reflected in the information-processing behaviour he exhibits, and in particular, the kind of information a learner seeks, the strategies he applies in learning situations, the kinds of information he generates from the environment, and the number of ways he can organize this information.

Cognitive styles is the term used to describe self-consistent, stable individual differences between learner's typical modes of cognitive organisation and cognitive functioning in the acquisition, processing and transmission of information. Cognitive styles are expected to reflect differences in learner's cognitive structures.

Cognitive preferences are one particular aspect of cognitive style which have been postulated as particularly relevant to the learning of information in schools. Cognitive preferences were originally defined by Heath (1964, p. 241) as "differing modes of attending to the subject matter of a course". Like cognitive styles, cognitive preferences are expected to reflect differences in learners' cognitive structures.

Three different conceptualisations of cognitive preferences

1. *Preferences for modes of attending to the subject matter of a course.* Heath (1964) hypothesised four cognitive preferences as particularly relevant for physics:

M – memory for specific facts and terms
A – practical application
Q – critical questioning of information
P – identification of a fundamental principle.

Subsequently this notion of cognitive preferences has been used as the basis of other cognitive preference tests in physics (Mackay 1971), chemistry (Marks 1967, Atwood 1967-8, King 1972, Kempa and Dubè 1973), biology (Barnett 1974), integrated science (Brown 1975), and in mathematics (Koay 1975).

2. *Preferences for modes of expressing mathematical ideas.* Travers, Heath and Cahen (1968) hypothesised three preferences for modes of mathematical expression:

V – verbal
S – symbollic
G – graphic

Subsequently Koay (1975) has developed a test in mathematics involving these modes.

3. *Willingness to respond consistently to various levels of knowledge.* Payette (1967) hypothesised preference for four levels of knowledge in the social sciences:

F – knowledge of specific facts
T – knowledge of terminology or terms
Me – knowledge of methodology
Ge – knowledge of principles and generalizations.

These levels are closely related to the *Taxonomy of Educational Objectives.*

Measuring cognitive preferences

One basic type of operation has been employed to measure all three types of cognitive

preferences.

The respondent is presented with series of items each consisting of a stem and a number of options, one for each of the cognitive preferences. The respondent is told that all options are factually correct, and asked to indicate his preferences for options in terms of their appeal, their relationship to the stem, or the way in which he/she would like material to be explained.

This form of measurement has one important problem – it yields ipsative scores, which precludes certain types of statistical analyses (Guilford 1967, Mackay 1971).

What is known about cognitive preferences?

Results of previous research (almost all on Heath's conceptualisation) indicate that:

- cognitive preferences change with time;
- end-of-course cognitive preferences differ with the nature of the curriculum;
- there is no evidence of any correlation between a teacher's cognitive preferences and the end-of-course cognitive preferences of his students;
- there is some evidence that changes in cognitive preferences are correlated with teaching practices they have experienced, although it has not been possible to identify the teaching practices concerned;
- cognitive preferences are related to performance on certain cognitive tasks. (For example, – physics students with a high preference for memory perform relatively better on tasks at the knowledge level in the Bloom *Taxonomy*; – physics students with a high cognitive preference for principles perform better on tasks at higher-than-knowledge levels in the *Taxonomy*; – physics students with a high cognitive preference for questioning perform relatively better on completion items; – physics students with a high cognitive preference for applications perform relatively better on multiple choice items) – see Figure 1.

All the above relationships are consistent with the hypothesised relationship between cognitive preferences, cognitive structure and performance on cognitive tasks.

An investigation of cognitive preferences and their relationship to some other variables.

The data pool. As part of the ETASC Project* three cognitive preference tests were administered to 481 servicemen's children in Grade 11 in all Australian states and territories. This was done to investigate the possibility that children who had experienced large numbers of school changes would adopt different learning strategies, in the form of different cognitive preferences, from those adopted by children who had changed school less frequently. In the event, there was no evidence that children who had experienced different numbers of changes of school had different cognitive preferences.

The data pool contained the following information on these children:

---

* An investigation of the effects of educational turbulence among Australian servicemen's children jointly sponsored by the Australian Departments of Education and Defence.

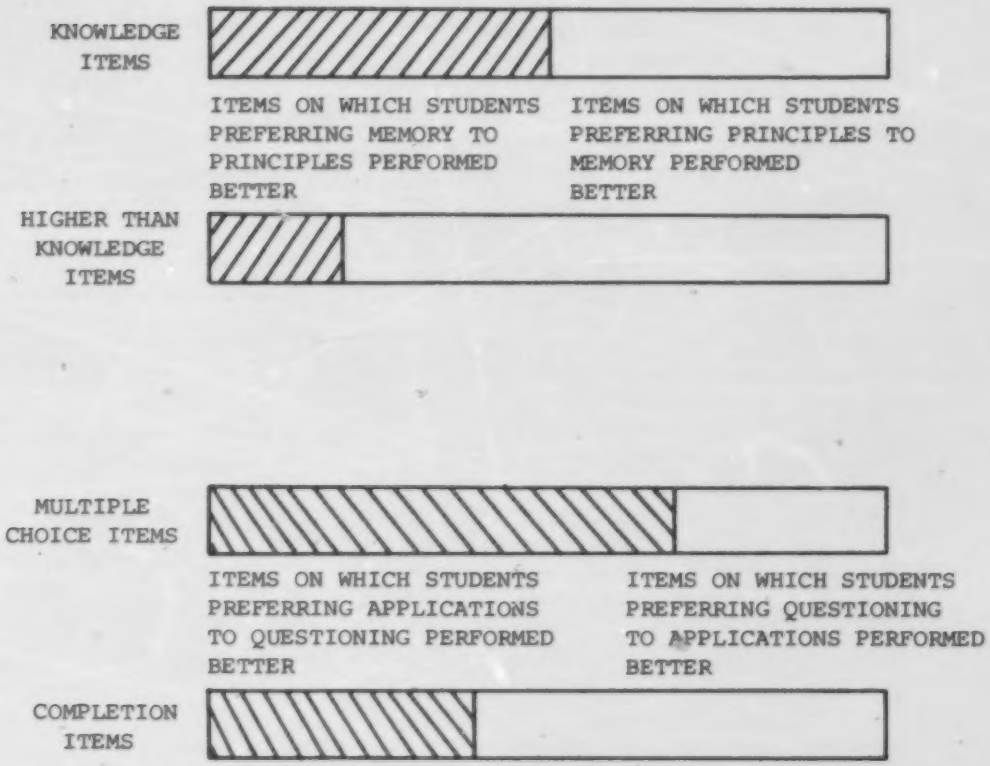

KNOWLEDGE ITEMS
ITEMS ON WHICH STUDENTS PREFERRING MEMORY TO PRINCIPLES PERFORMED BETTER
ITEMS ON WHICH STUDENTS PREFERRING PRINCIPLES TO MEMORY PERFORMED BETTER
HIGHER THAN KNOWLEDGE ITEMS
MULTIPLE CHOICE ITEMS
ITEMS ON WHICH STUDENTS PREFERRING APPLICATIONS TO QUESTIONING PERFORMED BETTER
ITEMS ON WHICH STUDENTS PREFERRING QUESTIONING TO APPLICATIONS PERFORMED BETTER
COMPLETION ITEMS

FIGURE 1: Representation of differences in performance of students with different cognitive preferences

1. *Eleven cognitive preference scores*

A Science cognitive preference test yielded four scores:

Preference M
Preference A
Preference Q
Preference P

} in Science (including items from biology, physics and chemistry)

A Mathematics cognitive preference test yielded three scores:

Preference V
Preference G
Preference S

} for form of expression of mathematical information

A Social Sciences cognitive preference test yielded four scores:

Preference F
Preference T
Preference Me
Preference Ge
} in Social Sciences

2. *General ability test scores* based on a revised and shortened version of the OTIS Higher Form B.

3. *Personality scores:*

Scores on two dimensions,

Extraversion-Introversion, and
Neuroticism-Stability

based on Eysneck's concept of personality (Eysenck and Eysenck, 1964).

4. *Socio-economic level of child's home environment*

This was based on father's Service rank which, in the Service context, is highly correlated with all the usual indices of socio-economic level.

5. *Sex of child*

The uses to which this data pool was put. Two basic questions were tackled using this data pool:

(i) What relationships exist between the eleven cognitive preference scores?
(ii) Are cognitive preferences related to –
general ability,
socio-economic level of home,
personality, or
sex of child?

In previous studies, cognitive preferences in different curriculum areas have been treated in isolation. In addition, whilst the relation between cognitive preferences and performance on cognitive tests has been explored, there has been little attempt to relate preferences to other more fundamental variables.

Relationships between the eleven cognitive preference scores. Two points are worth noting before the actual data are presented. Firstly, previous studies of cognitive preferences have been concerned with students' cognitive preferences in a curriculum area during a course of study in that curriculum area or at the end of a course of study in that curriculum area. These data were gathered for three major curriculum areas without regard to the actual course of study being undertaken by students. However, care was taken to ensure that students were likely to be familiar with all the subject-matter in the test. This was necessary to ensure that choices reflected preferences rather than familiarity with content.

Secondly, within each cognitive preference test, the sum of preference scores is equal to a constant, so that the scores in each subject are artificially negatively correlated. They are also ipsative in that they represent the respondent's relative preference among a number of stimuli rather than the absolute strength of his/her preference for any of the stimuli. For example, a very high preference score for Preference M in science corresponds to a preference for M in science relative to A, Q and P; it does not necessarily indicate a strong absolute preference for M at all.

The correlation between preference scores in different areas which are significant at the $\alpha = .001$ level are given in Table 1.

TABLE 1

Correlations between the eleven preference scores (those correlations which are significant beyond the $\alpha = .001$ level are included)

	M	A	Q	P	V	G	S	F	T	Me	Ge
M											
A											
Q								+.17			
P										+.16	
V											
G								+.21		-.19	
S											
F			+.17			+.21					
T											
Me				+.16		-.19					
Ge											

It is apparent from the results in Table 1 that:

- Cognitive preference for identification of underlying principles in Science is significantly positively correlated with cognitive preference for knowledge of methodology in Social Sciences.
- Cognitive preference for critical questioning of information in Science is significantly correlated with cognitive preference for knowledge of specific facts in Social Sciences.
- Cognitive preference for Graphical presentation of mathematical information is significantly positively correlated with cognitive preference for knowledge of specific facts in Social Sciences and significantly negatively correlated with cognitive preference for knowledge of methodology in Social Sciences.

All these significant correlations are consistent with the original conceptualisation of the cognitive preferences – however, the low correlations between preference scores in the various curriculum areas suggest that cognitive preferences in these areas are largely independent. Even after correction for attenuation, the largest correlation is only .40. (This is consistent with the finding of Atwood (1969) of no difference in Chemistry cognitive preferences of students studying different Physics courses.)

Subsequent factor analyses (after corrections for artificial negative correlations) indicate

relatively low overlap between cognitive preference scores on the three curriculum areas.

Relationships between cognitive preferences and other variables

The relationships between cognitive preferences and the other variables of:

general ability;
extraversion-introversion;
neuroticism-stability;
socio-economic level of home; and
sex of child

were investigated using a five-way ANOVA for each of the cognitive preferences in turn as the dependent variable. For this analysis, the population was divided into:

3 levels of general ability
3 levels of extraversion-introversion
3 levels of neuroticism-stability
2 levels of socio-economic level
2 levels of sex

Table 2 provides a summary of the main effects which were significant at the $\alpha = .05$ level.

TABLE 2

Summary of significant effects in ANOVA

	G.A.	Extroversion-introversion	Neuroticism-stability	S.E.L.	Sex
Preference M					
A	* (+)				
Q	* (-)		*		
P					
V					
G	* (-)				
S		* (+)			
F					
T	* (-)				
Me					
Ge					

These results are represented in Figure 2 and Figure 3.

There is evidence of a significant relationship between cognitive preferences and

general ability, and
personality,

but not evidence of a relationship with sex or socio-economic level. (Of course, socio-economic

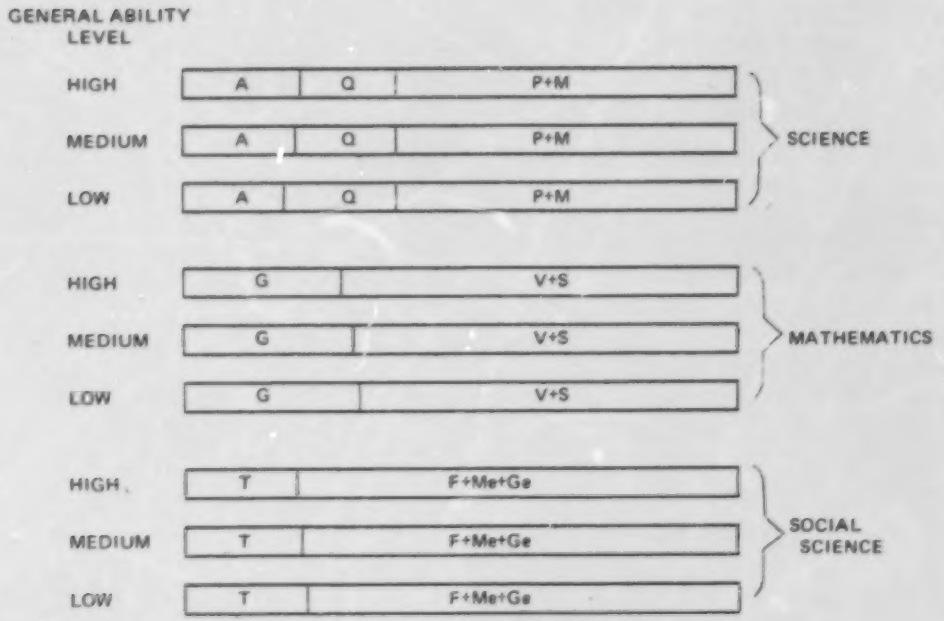


FIGURE 2: Differences in Cognitive Preferences of Students of Different Levels of General Ability

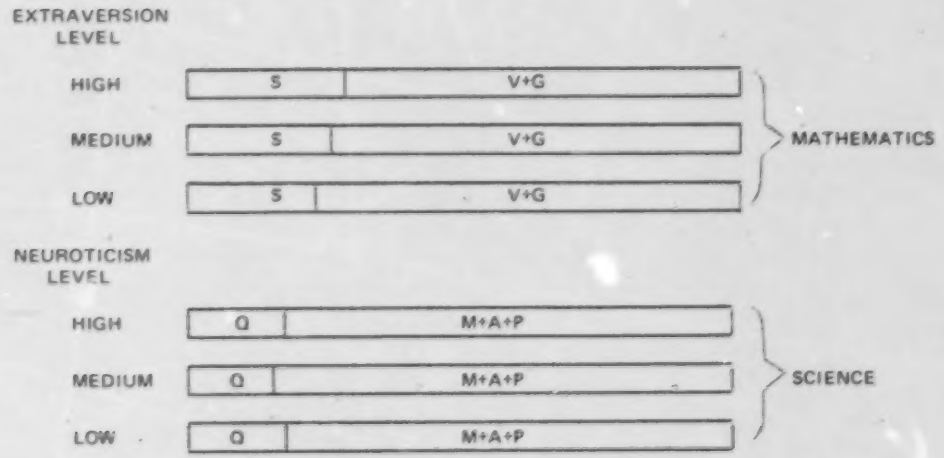


FIGURE 3: Difference in Cognitive Preferences of Students with Different Personalities

level and general ability are correlated (in this case r = .27), so that the socio-economic level main effect is after the effects of general ability are removed.)

Compared to students of lower general ability, students of high general ability showed:

- higher cognitive preferences for applications and lower cognitive preference for questioning in Science;
- lower cognitive preference for Graphical presentations of mathematical information; and
- lower cognitive preference for knowledge of terms in Social Science.

Compared to students who were more introverted, more extraverted students showed:

- higher cognitive preference for symbolic presentations of mathematical information.

Compared to students near the middle of this personality dimension, students who were more extreme in their neuroticism or stability showed:

- a higher cognitive preference for critical questioning of information in Science.

Summary and conclusions

1. Cognitive preferences in the different curriculum areas appear to be relatively uncorrelated, but in those cases where correlations occur, they are consistent with the hypothesised nature of the cognitive preferences.
2. Cognitive preferences are correlated with
   general ability, and
   personality,

   but there was no evidence of significant correlations between cognitive preferences and sex or socio-economic level (after the effects of general ability and personality are controlled).

References

ATWOOD, R.K. "CHEM Study achievement among groups classified by cognitive preference scores." *Journal of Research in Science Teaching*, **5**: 154-9, 1967-8.

ATWOOD, R.K. "Changes in cognitive preferences of high school physics students." *School Science and Mathematics*, **69**: 697-99, 1969.

AUSUBEL, D.P. *The Psychology of Meaningful Verbal Learning*. Grume and Stratten, New York, 1963.

BARNETT, H.C. "An investigation of relationships among biology achievement, perception of teacher style, and cognitive preferences." *Journal of Research in Science Teaching*, **11**: 141-7, 1974.

BROWN, S.A. "Cognitive preferences in science: their nature and analysis." *Studies in Science Education*, **2**: 43-65, 1975.

EYSENCK, H.J. & EYSENCK, G.B.G. *Manual of the Eysenck Personality Inventory*. University of London Press, London, 1964.

GUILFORD, J.P. "When not to factor analyse," in D.N. Jackson & S. Messick (eds.) *Problems in Human Assessment.* McGraw-Hill, New York, 309-17, 1967.

HEATH, R.W. "Curriculum, cognition and educational measurement." *Educational and Psychological Measurement,* **24**: 239-53, 1964.

KEMPA, R.F. & DUBÉ, G.E. "Cognitive preference orientation in students of chemistry." *British Journal of Educational Psychology,* **43**: 279-88, 1973.

KING, J.N. "Styles of thinking and curriculum comparison." M.Sc. Thesis, University of East Anglia, 1972.

KOAY, P.L. "Cognitive preferences and achievement in mathematics." M.Ed. Thesis, Monash University, 1975.

MACKAY, L.D. "Cognitive preferences and achievement in physics." Ph.D. Thesis, Monash University, 1971.

MARKS, R.L. "CBA high school chemistry and concept formation." *Journal of Chemical Education,* **44**: 471-4, 1967.

PAYETTE, R.F. *Development and Analysis of a Cognitive Preference Test in the Social Sciences.* University of Illinois, Illinois, 1967.

RIPPLE, R.E. "American cognitive studies: a review." *Journal of Research in Science Teaching,* **3**: 37-50, 1964.

TRAVERS, K.J., HEATH, R.W. & CAHEN, L.S. "Cognitive preferences in mathematics," paper read at A.E.R.A. meeting, Chicago, February 7-10th, 1968.

STUDENT PREFERENCES FOR YEAR II BIOLOGY CLASSES IN SOME SOUTH AUSTRALIAN SCHOOLS

Janet V. Keightley & Effice D. Best

Introduction

This paper reports the first stage of a project investigating some aspects of Year II Biology in South Australia schools. The aims of this part of the project are

(1) to provide feedback to teachers about the students' preferences for the Year II Biology course.
(2) to provide a general picture of teachers' aims for Year II Biology in South Australia.
(3) to provide a general picture of students' preferred areas of study in Year II Biology in South Australia.

Development of checklist

With the removal of the external examination at the end of the academic year and the existence of an advisory syllabus only, the checklists were developed so they would provide information useful to teachers in structuring their particular course to meet the bias of their individual classes. Three areas were considered when the checklist was being developed.

Best (1974) administered a student outcomes checklist to Matriculation students in South Australia and reported a factor analysis of these outcomes. The eight scales were labelled as follows'

a) Conceptual knowledge
b) Study skills
c) Scientific processes
d) Attitude change
e) Practical skills
f) "Familiarity breeds respect"
g) T.V./Newspaper
h) Factual knowledge.

These scales acted as one basis for the development of the checklists.

Convenient samples of students were given a short open-ended questionnaire at the end of their Year 10 studies to determine the kinds of activities and areas they were expecting to encounter in their study of Biology at the Year 11 level. The questions asked of them included

"If you intend to do Year 11 Biology next year,

a) What do you expect it to be like
b) What do you hope to gain from doing the course."

The answers to these questions were grouped and provided some of the examples given under the broad categories in the final checklist.

The course Objectives for Year 11 Biology classes using the *Web of Life* were also examined. (Teachers Guide *Web of Life*) The four overall aims of the course are to foster in the individual student

1) Scientific understanding of the living world, including man.
2) Understanding of the nature, scope and limitations of science.
3) Competence and willingness to apply biological understanding to scientific and everyday problems, and to approach these problems in scientific ways.
4) Ability to communicate effectively, and to work jointly with others towards solving problems of mutual concern.

The resultant checklist contained twelve broad categories, with examples to elucidate further the limits of the category. These are set out with the results in Table 1.

Procedure

At the beginning of the academic year, teachers in both state and private secondary schools were sent a copy of the student checklist and a plan outlining the nature and timing of the project. After being invited to participate a total of 129 teachers requested 3,500 copies of the student checklists.

Responses

Student. The students responded to the checklist items on a four point scale in terms of their preferences for learning about that particular category of Biology contained in the item. The scale points were "Very much like to learn", "Like to learn", "Don't care", and "Rather not learn". The students were also requested to choose the three items which they felt were the most important to them, but they were not asked to rank the three in order of priority.

The class data on student responses on preferred learning situations which was included in the second part of the checklist has been returned to the teachers involved and will not be reported in this paper.

Teacher. The teachers responded to the checklist items on a four-point scale in terms of their aims for the Year 11 Biology course. The scale points were "Major Aim", Minor Aim", "Incidental if at all", and "Not appropriate". The Teachers also were requested to choose the three items they regarded as being of highest priority. They too were not required to rank the three they had chosen.

The distribution of responses was as follows:

Number of schools	53
Number of classes	114
Number of students	2,616
Number of teachers completing teacher checklist	82

Frequency distributions for each item on both the student data and teacher data were calculated using Subprogram *Codebook* from the *Statistical Package for the Social Sciences*, (table 1).

Discussion of data

Teacher major aims. An examination of the frequency of teachers choosing each item as a major aim (Table 2) reveals a priority list of aims of a Year 11 Biology course which is closely allied to the stated objectives of the course in the Teachers Guide to the *Web of Life*. Items 1, 4, 11, 12, 10 and 7 are all stated in some form in the objectives and are consistent with the BSCS style of teaching which is an integral part of a course based on the *Web of Life*.

TABLE 1

Student and teacher answers to checklist

ITEM	STUDENT	%	%	TEACHER
1. **Understanding of the Biologists view of the world.** (For example: learning how animals, plants and man affect each other; how they survive; what kinds of problems biological knowledge can help to answer.)	Very much like to learn	23	94	Major Aim
	Like to learn	59	6	Minor Aim
	Don't care	15	0	Incidental if at all
	Rather not learn	3	0	Not appropriate
	No choice made	0	0	No choice made
2. **Knowledge of detailed facts about plants and animals.** (For example: learning the names of different parts of plants and animals and how they work.)	Very much like to learn	16	5	Major Aim
	Like to learn	46	59	Minor Aim
	Don't care	25	36	Incidental if at all
	Rather not learn	12	0	Not appropriate
	No choice made	1	0	No choice made
3. **Study skills.** (For example: learning how to use a library to find information; to think things out on your* own; to remember what you have learnt; to organize your work.)	Very much like to learn	22	59	Major Aim
	Like to learn	39	38	Minor Aim
	Don't care	28	5	Incidental if at all
	Rather not learn	10	0	Not appropriate
	No choice made	1	0	No choice made
4. **Scientific ways of thinking.** (For example: learning to observe; to analyse experiments to see what conclusions are justified; to think logically.)	Very much like to learn	21	87	Major Aim
	Like to learn	47	1	Minor Aim
	Don't care	24	1	Incidental if at all
	Rather not learn	7	0	Not appropriate
	No choice made	1	0	No choice made

* This item is cited from the student form. The teacher form was identical except for pronouns. (e.g. 'their' instead of 'your').

Table 1 continued

ITEM	STUDENT	%	%	TEACHER
5. **Interests and Attitudes.** (For example: gaining a greater interest in living things; greater appreciation of and respect for nature.)	Very much like to learn	32	80	Major Aim
	Like to learn	46	19	Minor Aim
	Don't care	18	1	Incidental if at all
	Rather not learn	3	0	Not appropriate
	No choice made	1	0	No choice made
6. **Practical skills.** (For example: learning to use a microscope; to dissect animals; to use laboratory apparatus.)	Very much like to learn	47	24	Major Aim
	Like to learn	32	69	Minor Aim
	Don't care	13	7	Incidental to all
	Rather not learn	6	0	Not appropriate
	No choice made	2	0	No choice made
7. **Biological knowledge with everyday applications.** (For example: learning how to choose foods for a healthy diet; how to make gardens grow better.)	Very much like to learn	34	26	Major Aim
	Like to learn	37	39	Minor Aim
	Don't care	22	35	Incidental if at all
	Rather not learn	6	0	Not appropriate
	No choice made	1	0	No choice made
8. **Basis for further learning.** (For example: gaining a background for Matriculation Biology.)	Very much like to learn	32	11	Major Aim
	Like to learn	34	43	Minor Aim
	Don't care	20	32	Incidental if at all
	Rather not learn	13	14	Not appropriate
	No choice made	1	0	No choice made

Table 1 continued

ITEM	STUDENT	%	%	TEACHER
9. Knowledge and skills useful for a job. (Write in examples you think applicable _____ _____ _____	Very much like to learn	38	11	Major Aim
	Like to learn	24	15	Minor Aim
	Don't care	11	48	Incidental if at all
	Rather not learn	8	26	Not appropriate
	No choice made	19	0	No choice made
10. Understanding of the nature of science. (For example: learning how scientists make discoveries; how scientists approach problems; how science has affected human history.)	Very much like to learn	11	33	Major Aim
	Like to learn	40	45	Minor Aim
	Don't care	37	20	Incidental if at all
	Rather not learn	11	2	Not appropriate
	No choice made	1	0	Not choice made
11. Communication. (For example: learning to describe observations; to write clearly; to be able to explain an idea to others.)	Very much like to learn	22	60	Major Aim
	Like to learn	45	35	Minor Aim
	Don't care	26	5	Incidental if at all
	Rather not learn	6	0	Not appropriate
	No choice made	1	0	No choice made
12. Co-operation. (For example: learning to work together with others on a problem; to see other people's point of view.)	Very much like to learn	29	48	Major Aim
	Like to learn	51	39	Minor Aim
	Don't care	17	12	Incidental if at all
	Rather not learn	2	1	Not appropriate
	No choice made	1	0	No choice made

While the acquisition of study skills is not specifically in the course objectives, in the majority of discussions on BSCS methodology the concept of the student being responsible for his/her own learning is stressed. The frequency of teachers choosing this as a major aim is not inconsistent with the overall goals of the courses.

TABLE 2

Percentage of teachers regarding items as a major aim

Percentage	Item
94	1. Understanding of the biologists view of the world.
87	4. Scientific ways of thinking.
80	5. Interests and Attitudes.
60	11. Communication.
59	3. Study Skills.
48	12. Co-operation.
33	10. Understanding the nature of science.
26	7. Biological knowledge with everyday applications.
24	6. Practical skills.
11	8. Basis for further learning.
11	9. Knowledge and skills useful for a job.
5	2. Knowledge and detailed facts about plants and animals.

There were some inconsistencies between the percentage of teachers regarding an item as being a major aim and the aims of the *Web of Life* course. One was the relatively low percentage of teachers choosing the *Web of Life* major aim of "understanding the nature of science". While teachers certainly do not rank this aim as unimportant, they do see it as being less important then those of study skills, and interests and attitudes. These, although not stated in the course objectives, were held in higher regard by more teachers.

Student preferences. Table 3 reveals that there is very much less consensus among the students population examined than among their teachers. The diversity of student interest, bias and expectations is reflected in these data. Among the students responding there were 13 percent who would rather not learn biology as a basis for further learning; however, there were 32 percent who would very much like to learn biology for this reason. Similar divisions in responses were given for item 2 (knowledge of detailed facts about plants and animals) and item 3 (study skills).

The student data on item 9 reveals that 19 percent of students chose not to answer this particular item. This suggests that students either do not see biology as being relevant to any job in particular or else they chose not to answer a question where they were also asked to provide an example of the particular job they had in mind.

TABLE 3

Percentage of students regarding items as very desirable

Percentage regarding item as very desirable	Item
47	6. Practical skills.
38	9. Knowledge and skills useful for a job.
34	7. Biological knowledge with everyday applications.
32	5. Interests and Attitudes.
32	8. Basis for further learning.
29	12. Co-operation.
23	1. Understanding of Biologists view of the world.
22	3. Study skills.
22	11. Communication.
21	4. Scientific ways of thinking.
16	2. Knowledge of detailed facts about plants and animals.
11	10. Understanding of the nature of science.

Comparison of teacher and student priorities. The data in Table 4 represent the percentages of teachers and students who recorded the particular item as being one of the three most important in the checklist. These percentages were then ranked to give an indication of the order of priority of the items with respect to each group.

While the teachers ranked an understanding of the biologists view of the world, interests and attitudes, and scientific ways of thinking as being the three highest priority items these items only ranked fourth, fifth, and eleventh on the student priorities. An even greater discrepancy is revealed when the teachers' ranking of the three highest ranking items on the student responses is examined. Practical skills, knowledge and skills useful for a job, and biological knowledge with everyday applications rank only equal tenth and seventh on the teachers rankings.

Implications

If students have chosen the study of biology because they expect it to cover the areas they would very much like to learn, it would appear that teachers have not successfully conveyed their aims to the students. However, some students may have chosen to continue the course in a hope that some of their preferred areas of study would be covered to some degree.

The discrepancy between what teachers hope to teach and what the students hope to learn suggest that the students are more instrumental in their approach to biology than the teachers. The teachers, on the other hand are more conceptual in their approach.

The teacher can change the aims of the Year 11 Biology course in the light of the student preferences. However, if such a change is not thought to be desirable, a teacher can catch student

TABLE 4

Teacher and student priorities

Item	Percent including item in top three priorities		Rank	
	Teachers	Students	Teachers	Students
1. Understanding of the biologist's view of the world.	74	29	1	4
2. Knowledge of detailed facts about plants and animals.	1	18	12	9
3. Study skills	20	17	6	10
4. Scientific ways of thinking.	55	14	3	11
5. Interests and attitudes.	65	27	2	5
6. Practical skills.	2	38	10	1
7. Biological knowledge with everyday applications	14	30	7	2
8. Basis for further learning.	5	24	9	7
9. Knowledge and skills useful for a job.	2	30	10	2
10. Understanding of the nature of science.	6	11	8	12
11. Communication.	26	19	5	8
12. Co-operation.	27	26	4	6

interest via practical activities and applications. The teacher can use such practical activities and applications to lead the students to the concepts.

If teachers want to tailor their teaching to the student preferences, the lack of consensus of opinion in the student population indicates that individual goals must be set for students. This is in addition to the notion of individual progression.

References

BEST, E.D. "Student perceptions of activities and outcomes in South Australian Matriculation Biology Classes." *Research in Science Education*, 4: 151-159, 1974.

MORGAN, D. (ed.) *Biological Science: The Web of Life, Teachers' Guide, Part 1.* Australian Academy of Science, Canberra, 6-8, 1973.

AN EVALUATION OF THE INTRODUCTORY CLINICAL SKILLS COMPONENT OF THE 3RD YEAR MEDICAL COURSE AT ADELAIDE UNIVERSITY

W.C. Hall & J. Lawrence

Introduction

At the invitation of the Community Medicine: History Taking Sub-Committee, the Advisory Centre for University Education (ACUE) was invited to evaluate weeks 28 and 29 of the 3rd year program. The evaluation was directed at the effect of the course on students and was not an assessment of either lecturers or tutors.

One hundred and twenty two 3rd year medical students attended the two weeks. These students were listed in descending order of their aggregate mark in their first professional examination (3rd year). Odd numbered students were labelled "A students" (sA) and even numbered students were labelled "B students" (sB). Different evaluations were carried out by sA and sB.

Thirteen students (6 sA and 7 sB) falling within the 40-60 percentile range of the professional examination were given sound cassettes for use during one interviewing session and one clinical examining session. The progress of two volunteers (s'A and s'B) obtained from the 13 students was followed fairly closely and, in particular, they were televised on two occasions.

A member of the ACUE attended some of the sessions as an observer and held informal discussions with students and staff.

The main purpose of the investigation was to determine how successfully students and staff believed the five course aims, which were defined before the course, were achieved. This was attempted by means of questionnaires, free response written accounts and direct observation (including video and sound recordings).

An interesting feature of the evaluation is the progressive narrowing down of the study from the whole class to two individual students.

Background to the study

In the traditional medical course, training in basic sciences – anatomy, physiology, biochemistry etc. – occurs in academic departments in a typical university environment and is largely supervised by tenured university staff.

Clinical training is a mixture of traditional teaching of subjects (such as microbiology) and training in clinical skills and knowledge. Much of the clinical training takes place in hospital wards. Teachers are usually practising doctors over a wide range of seniority with relatively loose University affiliation. The two basic clinical skills are the clinical interview and the physical examination leading to appropriate further investigations, diagnosis and therapy. Much of the teaching is located at the bedside and it is essential that students develop clinical skills by interviewing and examining patients.

Traditionally much emphasis has been placed on learning by doing and intuition with limited attempts to direct and train students in the psychodynamics of the clinical interview.

Major problems in this investigation in devising an effective introduction to learning the new skills included:

(a) adaptation of student to the rather rigid new hospital environment with its

hierarchies of doctors, sisters and other health professionals;

(b) natural insecurity of students and shyness in approaching patients;

(c) doubt about the role of junior students in the hospital ward;

(d) lack of knowledge of clinical medicine, pathology, etc. so that the relevance of details of history and examination is unclear. The consequent insecurity makes students anxious about communicating with patients;

(e) the belief that the most effective learning occurs in small groups using tutorial style. This required a large number of teachers, most of them relatively enthusiastic but untrained in teaching techniques and fitting teaching into their professional responsibilities;

(f) the traditional emphasis on the individuality of communication techniques between doctor and patient and thus lack of concensus among teachers concerning techniques of the clinical interview.

In the Adelaide Medical School the transition from "pre clinical" medicine takes place at the beginning of the 3rd term of 3rd year. At this time clinical teachers are also teaching fourth year students and so the supply of potential tutors is somewhat limited. The introductory week which preceded the weeks of the study included small group tours of various hospitals to familiarise students with the new environment and to encourage them to talk to staff and patients. During the next two weeks students had a series of lectures, large group discussions, brief films and demonstrations occupying 20 hours; tutorials in hospital wards with encouragement to seek and talk to hospital patients were scheduled for at least 14 hours. Students were encouraged to use tape recorders and to attempt group learning. In addition to the stated aims a further objective of the course was to introduce some comprehension of the logic of the clinical diagnostic process and decision making.

The tutors were medical graduates of various ages and seniority usually less than 35 years old. They were not briefed at group sessions.

The need for curriculum evaluation

Curriculum evaluation is frequently discussed but rarely practiced! It is a messy business, not often permitting the experimental designs summarised by Campbell & Stanley (1963) but being "quasi-experimental" in its approach. As described by Hall (1975) what is usually being attempted is a description (quantified, if possible) of what is actually occurring, followed by appropriate feedback so that modifications can take place after comparisons have been made with the expected outcomes.

The earlier work of Freeman & Byrne (1974) had been noted although their study was largely concerned with course assessment rather than curriculum evaluation.

Course aims

During the two weeks it was hoped that at least five aims would be achieved by students in their introductory clinical course. These are listed below.

A. Students should know the basic principles and techniques of clinical interview and physical examination.

B. Students should know the more common problems which can arise as a result of interviewing and examining a patient and know how to deal with these problems.

C. Students should be able to write up a case history in a problem-orientated manner.

D. Students should know what the relationship between a doctor and a patient should be.

E. Students should not feel anxious in the presence of patients, doctors and nurses and when in a ward and in the hospital.

All members of staff had this list of aims in their possession before the start of the two weeks. Students were told the aims on the Friday preceding the two weeks and sA were provided with a printed list as part of the introductory notes to their evaluation sheets.

The study

Evaluation by staff. Staff were given an evaluation form which consisted of course aims, examples of criteria to be used in assessing their contribution to the achievement of these aims, and an evaluation matrix. The attempt to specify criteria is worth a mention. Rating scales are notoriously unreliable and merely to ask for an 'excellent', 'good' etc. grading is generally a waste of time. By way of example, the criteria used in evaluating the achievement of Aim E is shown below:

	1 Of very great importance	2 Of great importance	3 Of some importance	4 Of no importance
AIM E	Recognised that there are anxieties to overcome and gave detailed explanations of ways of reducing these	Explained in detail ways of reducing anxiety in at least one of the five areas (patients doctors, nurses, ward, hospital.)	Briefly mentioned ways of reducing anxiety.	Did not cover any aspect of this aim.

The matrix completed by staff is shown below:

RATING

AIM	1	2	3	4
A				
B				
C				
D				
E				

Evaluation by students. sA were given an evaluation form similar to that completed by staff except

that there was a matrix to complete for every session during the two weeks.

sB were asked to evaluate just one of the course aims, Aim E. Instructions for completing their form are reproduced below:

> One of the aims of the two weeks' course from 9-20 September is
>
> "Students should not feel anxious in the presence of patients, doctors and nurses and when in a ward and in the hospital."
>
> Those involved in teaching this course would like to know whether your anxieties of the five variables (patients, doctors, nurses and the ward and hospital environments) alter during the two weeks and, if so, how.
> Evaluation sheets for each of the two weeks for all five variables are attached. After each session you are asked to indicate how your anxiety has altered. Everyone starts at 0 on Friday, 6th September. After each session –
>
> if anxiety has increased slightly, mark 1 number further to the left
> if anxiety has increased greatly, mark 2 numbers further to the left
> if anxiety has decreased slightly, mark 1 number further to the right
> if anxiety has decreased greatly, mark 2 numbers further to the right
>
> For example, if on one of the variables anxiety has decreased greatly during the 4.00 p.m. session on Monday, 9th September, you might show this change as follows:
>
> 2.00 ——|——————————————————————
> -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10
>
> 4.00 ————|————————————————————
> -10 -9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10
>
> If you reach -10 or +10 on a scale, then please start from 0 again.

All students were given an open-ended questionnaire asking for comments on the course (including successful features and improvements which could be made). These were distributed at the end of the course.

S'A and s'B were provided with a short questionnaire for use during their viewing of the video-recordings.

Sound cassettes were supplied to 13 students for later evaluation. Video-recordings were made of s'A on Friday 13 September and of s'B on Friday 20 September.

An observer (Dr. W. Hall) kept notes during the two weeks, particularly of comments made by s'A and s'B.

The possibility of devising questionnaires to be completed by patients was discussed but, because of lack of time, was not pursued. It would be worth considering this at an earlier stage of any future evaluation.

Results

Questionnaire A. Seventy students completed the survey (i.e. 58%). These consisted of 33 sA (including s'A) and 37 sB (including s'B).

The mean and standard deviation for each assessment of all five aims by sA were calculated.

A comparison was done of the evaluations made by sA (mean value), s'A and staff. In making the comparisons, differences of 1 (or greater than 1) were considered to be 'significant'. For differences less than 1, agreement was assumed to be the same.

It was possible to identify those sessions where the students' estimates of the achievement of course aims were more encouraging than staff estimates and *vice versa*.

On the basis of the ratings it was suggested that there was probably low achievement on

three of the aims (B, C and D).

The interesting features of the evaluations by sA, s'A and staff were:

(a) an attempt was made to define criteria to be used in making the evaluations;
(b) the progress of one 'average' student was carefully followed and his evaluations compared with those made by the larger sub-group sA;
(c) staff used exactly the same questionnaire as that completed by students.

Questionnaire F. One of the assumptions made before the course was that students would probably be anxious about their contacts with patients, doctors and nurses and when in a ward and in the hospital (Aim E). An 'anxiety scale' was devised for each of these five variables and the mean positions on this scale and the standard deviations were calculated. The results for the sub-group sB were compared with those by s'B. The results which were obtained for sB are plotted in Figure 1.

Trends in changes of anxiety were being sought; the scale is not absolute and it is meaningless to look at results for a single session.

This technique for plotting gross changes (i.e. "macro-evaluation" instead of "micro-evaluation") is currently being repeated for another course evaluation. A 3 point scale only is being used (+1, 0, -1) and it is likely that this will continue to be the approach in future studies of this nature.

Computations were performed on the final results obtained on the anxiety scale and for all five variables they were significant at the 0.05 level.

All five variables show similar trends:

(a) Slight increase in anxiety (four variable) during day 1.
(b) Decrease in anxiety for days 2 – 4.
(c) Increase in anxiety for days 5 – 6.
(d) Decrease in anxiety for days 7 – 10.

The physical examination of patients took place only in week 2 and so this could explain the changes in anxiety. It would be interesting to repeat the evaluation with the examination taking place at the start of the course.

The interesting features of the evaluations by sB and s'B were:

(a) the progress of one 'average' student was carefully followed and his evaluations compared with those made by the larger sub-group sB;
(b) the attempt to devise a scale for plotting gross changes (in this case 'anxiety') over a period of time.

Free-response questionnaire. Thirty two students (26%) completed this questionnaire. The classification of replies showed (among other points) that although students were enthusiastic about their contact with patients, they were critical of the small amount of time actually spent in their presence. They were also critical of the lack of coordination between hospital tutors and university staff and of the large tutorial groups.

Detailed comments were also made by s'A and s'B on particular aspects of the course. In his evaluation of Aims A-E, s'A was more critical than the sub-group from which he was drawn, which had to be borne in mind when his detailed comments were considered.

The students, by and large, identified those areas as problems which we had expected. In

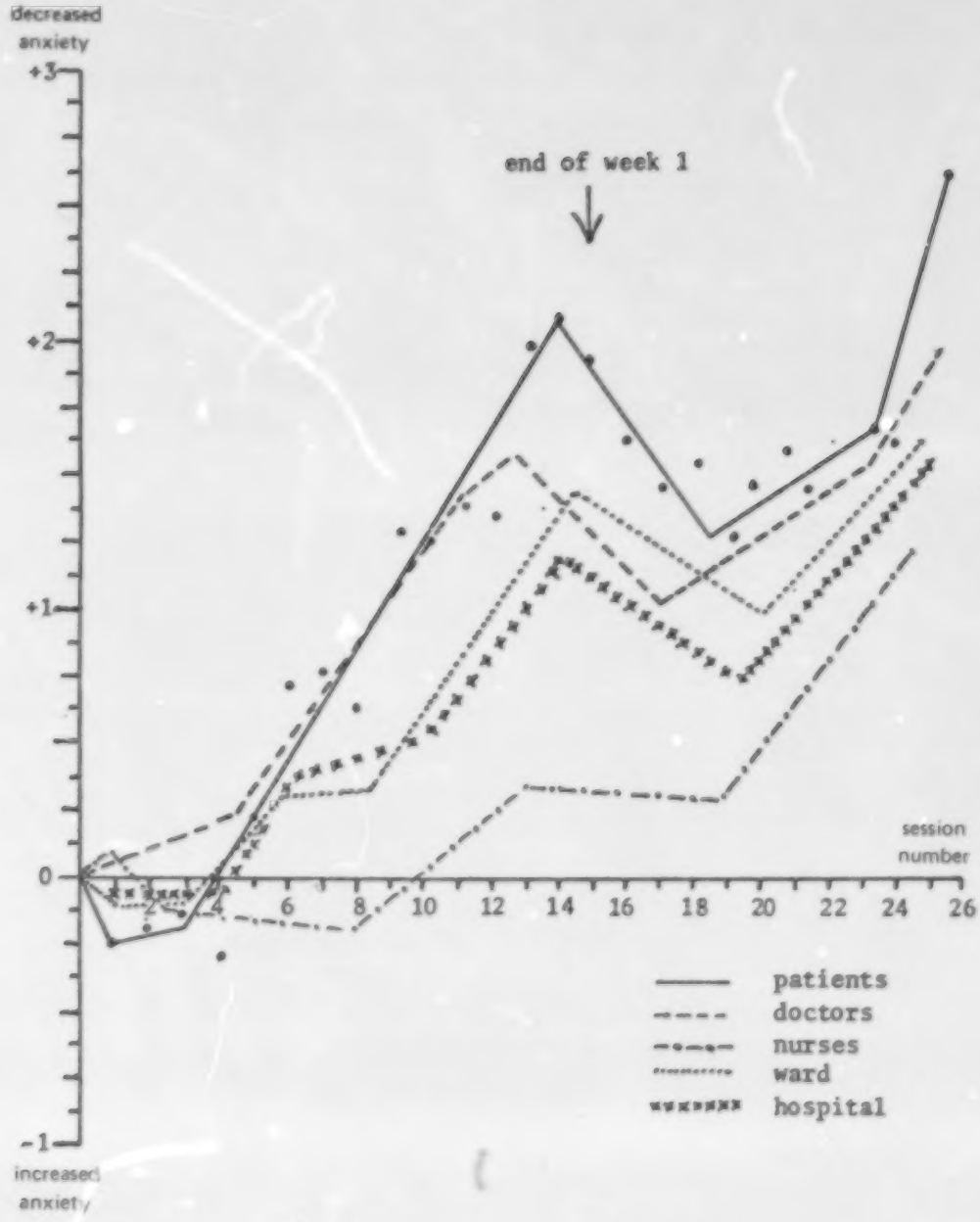


FIGURE 1: Changes in Anxiety.
(The solid dots are the exact plots for the 'patient' variable.)

their enthusiasm for patient contact they resented the large number of tutors who talked at them rather than facilitating the student "learning for himself" by experience with patients. This arose largely because the teachers were not adequately coordinated and directed or trained enough so that

(a) they were not fully aware of the lecture content and principles in spite of having student lecture notes and suggested texts distributed to them;
(b) they did not follow the attempt to program the introductory tutorials;
(c) tutors did not support the suggestion made to students that they utilize group interaction with tape recordings etc.

These are also possible explanations for the low achievement of three of the aims.

Recordings. Sound cassette and video recordings were used for student self assessment purposes and also to evaluate student performance in achieving the course aims.

Both s'A and s'B were pleased to see that their performances were not as bad as they had imagined them to be. In particular, although they had both felt nervous during the recordings, the play-backs revealed little nervousness and so this helped to increase their confidence.

Possible changes to the course as a result of the evaluation

One outcome of the discussions which took place after the evaluation was the realisation that too many aims were attempted. It is now recognised that in two weeks it is impossible to achieve much more than Aim A (which was, in fact, satisfactorily accomplished).

Secondly, the coordination between the work of the hospital tutor and the university teacher needs to be strengthened. Lectures and hospital sessions are supposed to fuse.

Thirdly, immediate feedback to students on their performances in the hospital should be timetabled: the changes in "fear" could be quite different if this is done.

Finally, the degree of involvement of one student (s'B) was so intense that he produced his own proposals for a two weeks course as a result of the evaluation!

References

CAMPBELL, D.T. & STANLEY, J.D. "Experimental and quasi-experimental designs for research on teaching," in Gage, N.L. (ed.) *Handbook of Research on Teaching*. Rand McNally, Chicago, 1963, 171-246.

HALL, W.C. (1975). "Evalutating university courses", *Australian University* (in press).

FREEMAN, J. & BYRNE, P.S. *The Assessment of Postgraduate Training in General Practice*. Society for Research into Higher Education, London, 1974.

AN EXAMPLE OF THE ROLE OF BIOMETRICS IN TEACHING PRACTICAL ENVIRONMENTAL SCIENCE

P.R.B. Blood & A.H.A. Bensink

Introduction

Tertiary science courses with ecological and environmental emphases have proliferated in the last half-decade. Innovative degree courses in integrative areas such as "environmental studies" and "natural resources" are being offered by many institutions, permitting the more classical courses to be retained in comparatively intact form. Although many of these classical courses have also been influenced by the environmentalist lobby, some still fail to examine adequately environmental issues and their social and ethical implications. Agricultural Science is perhaps one such course.

Traditionally, tertiary agricultural science education has been concerned with imparting the wide spectrum of physical, biological and socio-economic bases and skills necessary for the efficient production of crops, livestock etc. On the whole, scant attention has been paid, in the past, to the environmental impact of agricultural practices and little effort has been invested in utilising the many resources, physical and mental, available to agricultural scientists to resolve or mitigate the undesirable consequences of modern agricultural practices. As Norman (1973) succinctly puts it, agricultural science students have "the specific feeling that at present agricultural scientists are cast in the role of the blinkered technolgist who cannot appreciate the ecological consequences of his recommendations".

Recently, Hill and Smith (1973) investigated the degree to which the aims of an "ideal" final year agricultural science course were extant in existing ("actual") courses at the University of Melbourne and La Trobe University. Aims commonly accepted for environmental education and individual development, and an aim basic for courses in agricultural science, were included. There was congruence between "actual" and "ideal" courses for the agricultural science aim but not for aims concerned with environmental education and individual development.

Hill and Smith's study indicates the areas in which modern agricultural science courses are failing to satisfy student needs. The same dissatisfactions were discussed in depth by Davies (1970) in his excellent essay on the "crisis in biology". Both emphasize the need for curricula reform and formative evaluation studies in this subject. Reinforcement of this principle was provided by Atkin (1968) and Tisher (1971) but does not appear to enjoy the currency it should demand.

This paper describes initial stages in the development of a tertiary Agricultural Science and Pure Science course offered by the University of Queensland. Unexpected outcomes of the teaching approach are discussed and possible causal factors of these outcomes are pinpointed for quantitative analysis and special consideration in the development of a programmed teaching strategy.

Background

One aspect of agricultural science that has closely impinged on social consciousness is crop protection. The large-scale indiscriminate use of pesticides during the last 30 years has caused so many problems and side-effects that pubric outrage has effected the outright banning or restriction of many otherwise useful chemical protectants. This is a trend which does not appear to be diminishing, especially in the more developed countries of the world (Smith, 1975).

Training in the area of crop protection during the last 20 years has centered on the routine testing of pesticides and the technology and methodology of their application. Decision-making in crop protection became merely a function of comparing candidate compounds for efficacy and cost. This simple and empirical approach was reflected in the tertiary courses designed to provide professional training in crop protection and pest control.

The emergence over the last 15 years of a number of undesirable phenomena associated with chemical control has prompted a world-wide review of the desirability of placing undue reliance on unilateral chemical control approaches. Some of these phenomena are recognized by society at large, e.g. chemical pollution caused by pesticides and their biological concentration in higher trophic levels leading to wildlife degradation. The chemicals implicated here belong to the persistent, long-term degradable chlorinated hydrocarbon group, e.g., DDT. Another group of compounds, the organo-phosphates, not as persistent in the environment as the former group, are highly toxic and present a considerable public health hazard.

There are other phenomena of which professional crop protectionists and ecologists are aware but which are not known to the general community. The first is a genetic phenomenon, the evolution by selection of pesticide resistance in pests, especially insects and related arthropods. There are now several hundred examples throughout the world of insect pests resistant to one or more insecticides. The natural reaction to the build-up of resistance has been more freuqent application of more concentrated or substitute compounds. This reaction has inevitably resulted in an exacerbation of the whole situation.

Insecticides invariably discriminate against beneficial arthropods in agroecosystems. This has led to the phenomenon known as "flare-back", the result of an insecticide-mediated release of the pest from its natural enemies (predators and parasitoids) which are slower in building up to effective densities post-application. Yet another phenomenon is the gradual emergence into economic prominence of hitherto "pest-neutral" insects. Again, this is a result of insecticides controlling the primary pest but concomitantly destroying the natural enemies of the secondary pest. The ultimate phenomenon is the increasing economic pressure generated by the failure of unilateral chemical approaches to crop protection. The inevitable consequence of this cost-squeeze has been the syndrome of resorting to an ever-increasing reliance on new compounds, the reluctance to gamble on multi-lateral approaches and the eventual switch to alternative crop systems possessing less refractory protection problems.

The undesirable impact of pesticides on agroecosystems has led to the development of multilateral approaches in solving crop protection problems. The terminology used to describe this methodology has variously and to some extent erroneously included, "applied natural control", "biological control", "ecological control" and "integrated control". Most workers in the area prefer the term Integrated Pest Management (IPM) which may be defined as a management system based on the utilization of intrinsic natural controls (i.e. environmental resistance) and all suitable extrinsic controls harmonised into a program based on rational ecological and socio-economic considerations (Smith, 1975).

Geier and Springett (1974) categorize three broad approaches to pest control; the opportunistic which is concerned with immediate objectives wherein the methodologies employed are judged solely on their technical effectiveness; the gradualistic which is based on the tenet that improvement accrues in the step-wise manner of continual hypothesis-erection and validation; the holistic which demands a radical departure from current conceptual attitudes to achieve goals in a short, critical span of time using a total-systems approach.

The last-named approach is operational to date only in North America. Most developed

nations are striving to emulate American successes but the vast majority of management programs in these countries lie somewhere between the opportunistic and the gradualistic approach. The current situation in Australia, as elsewhere, is most aptly described by quoting from Geier and Springett who warn that ". . ., it could be dangerous to let pest control proceed haphazardly as it does now at most times and places i.e. under ad hoc leadership guided by part-time experts, narrow specialists, or advocates for special interests, whose decisions are implemented mostly by untrained operatives". The message here is particularly apposite to the changes needed in crop protection vocational education. There is a pressing need to train future operatives in the holistic or IPM approach for a number of reasons, not the least of which is the potential infusion of new ideas in problem-solving into the decision-making infra structure of political and commercial institutions, (Blood, *et al.* 1975).

IPM training in Queensland

The University of Queensland transferred to a semester system in 1974 and this permitted considerable curriculum revision and the proliferation of smaller courses adapted to semesterisation. The events of 1974 allowed the compartmentalising of a course "Applied Entomology", given by the Department of Entomology prior to 1974, into three derivatives: "Economic Entomology", "Insecticidal Control" and "Pest Management". The course objectives of the last mentioned are to:

1. Survey the full range of non-insecticidal control technologies and the methodologies of their implementation.
2. Define the role of insecticides in multi-lateral pest management programs.
3. Elucidate the mechanisms involved and measures required in developing rational pest management programs.
4. Examine incorporation of pest management programs into the higher hierarchies of farm management and regional economic management.

The course was organised into 26 one hour lecture periods and 13 three hour practical periods. Both the lecture and practical courses were designed with a decreasing emphasis on course objectives 1 through to 4. This bias in priorities allocation arose for a number of reasons; firstly, the inchoative nature of the course demanded retention of most of the elements of the antecedent parent course; secondly, there was a vocational obligation to provide full exposure to all alternative technologies; thirdly, the mixed backgrounds of the students would presumably result in individualistic responses to the issues generated by objectives 3 and 4; fourthly, as the course was "breaking new ground" to a certain extent in objective 3 and to a large extent in objective 4, it was considered important to design course '74 as a trial which would, upon evaluation, indicate more appropriate strategies for achievement of the desired course objectives in subsequent years.

Thus a course sullabus resulted wherein lecture periods were almost fully occupied with the transferral of the burgeoning material implicit in course objectives 1 and 2 and to a certain extent 3 and 4. The practical component was designed to reinforce all course objectives, but particularly 3 and 4, by recourse to a field experiment chosen to simulate a number of management options open to a commerical grower. Our assumption was that, by allowing students an integral role in the design, execution and analysis of the field experiment, each would individually formulate a realistic evaluation of management options available to the farmer-entrepreneur in the context of current pronouncements on environmental issues.

Although an educational device such as this may be criticised as too *recherche*, the dearth of local research which can be translated into instructive case histories at present precludes the adoption of any other scheme. Laboratory investigation of some components of an IPM program could be considered as disjointed and unrealistic, while a computer simulation and discussion, whilst instructive, could also be criticised for lack of realism.

The IPM field experiment

Our solution, then, was to create an agroecosystem on campus and to impose upon it a series of IPM "treatments" so that each student could individually evaluate their effects on pest populations and crop yields and so develop an IPM strategy for the target system.

Cabbages were chosen as the crop for the object agroecosystem because they possess distinct advantages for this exercise:

1. They can be grown in the Brisbane metropolitan area thus obviating the need to travel long distances to field plots.
2. They are a winter crop requiring approximately 15 weeks growth and thus compatible with the academic year and the semester period.
3. Monitoring of pest populations associated with the crop can be accomplished by non-destructive sampling techniques.
4. They are a crop requiring no special agronomic treatments and the application of IPM treatments poses no serious difficulties.
5. The assessment of marketability in terms of yield and cosmetic appearance is rapid and uncomplicated.

The experimental area on campus was cultivated, fertilised and generally treated in line with commerical grower and Department of Primary Industry recommendations. The experiment was laid out as a randomized complete block design with 3 replicate blocks. Each block comprising 27 plots was surrounded by guard rows to minimise "edge effects". Cabbages were planted 12 to a plot. The 27 treatment combinations were assigned at random to plots within each block.

The IPM treatments. Of the many IPM options available to the pest manager only a few were chosen in this experiment because of restrictions in resources of time, personnel and finance. We were constrained to the employment of 3 basic options, namely, varietal, cultural and insecticidal control.

Varietal control is the utilisation of varieties or cultivars of crops which exhibit resistance to the effects of or to pest agents themselves. This resistance may manifest in three ways:– (1) Non-preference – in which the plant possesses factors that render it unattractive to insect pests for their oviposition, feeding or shelter. (2) Antibiosis – where the host plants adversely affect the bionomics of the pests feeding on them. (3) Tolerance – in which the plant suffers minimal damage in spite of supporting pest populations adequate to severely damage susceptible hosts.

The experiment tested 3 common, commercially acceptable cabbage varieties.

Cultural and management procedures for any agroecosystem must firstly keep the crop healthy by appropriate agronomic practices. Cultivation procedures can be manipulated in various ways and one approach receiving considerable attention is the growth of certain crops in the presence of naturally occurring weeds or a single plant species. The rationale here is to enhance the attractiveness of the system to predators and parasites of the major pests by providing shelter, resting sites, pollen and nectar supplies upon which dipterous and hymenopterous parasites mature

their eggs or nutritionally sustain themselves.

The 3 cultural treatments imposed consisted of (1) no treatment i.e. the natural community of weeds was allowed to grow unhindered, (2) a cover crop, red clover – a suitable plant for pollen and nectar provision, (3) provision of black plastic which prevented emergence of all weeds.

There is virtually complete reliance on chemical insecticides for the control of pests on cabbages grown commercially in Queensland. Currently the Queensland Department of Primary Industry recommends the use of an organophosphate compound Diazinon (0,0-diethyl - 0 - (2 - isopropyl - 4 - methyl - 6 - pyrimidinyl) phosporothioate). This compound is generally non-phytotoxic, is a safe insecticide from the point of consumer health and tends not to accumulate in the environment and in food chains but is extremely toxic to natural biocontrol agents. One of the insecticidal treatments utilised this compound.

The second treatment employed "Dipel" a trade-name for a microbial insecticide composed of a bacterium *Bacillus thuringiensis* (B.t.) and its associated crystalline endotoxin. This microbe has the effect of inhibiting feeding of the pest and results in death of the pest after several days. The bacterium is specific to the pest and does not affect biocontrol organisms in any way, is non-pollutive and is non-hazardous to operators.

Instruction and information inputs. Instructions and information issued to students throughout the entire experiment were as follows:

1. An introductory lecture supplemented by a background information brief, in which the concept of IPM was discussed and the objectives and design of the experiment were detailed.
2. Field instruction on plot layout, random allocation of treatment combinations, cultural and agronomic procedures. Students planted the crop and assisted with the cultural treatments.
3. A discussion/instruction session concerning possible approaches to monitoring the effects of the treatments on pests, predator and parasite populations. Following this, the chosen methods were detailed in an instruction hseet and followed throughout the experiment. Six plants/plot were monitored weekly by students, individuals being assigned different plots each week.
4. An instruction session on the identification of each life history stage of the major cabbage pests supplemented by a list of insects (pests, predators and parasites) and pathogens commonly associated with the crop.
5. A discussion on approaches to the analysis and interpretation of results and instruction on the role and use of ANOVA and the Least Significant Difference (L.S.D.) *a priori* test. Students were taught to feed their pre-treated data into appropriate computer programs via a remote terminal. The results of basic statistical analyses were distributed to each student. There was also ready access to desk calculators and the computer terminal for those wishing to perform additional analyses.

Other information distributed to students included:

1. Data on total leaf surface area for each cabbage variety at various stages of crop growth.
2. Data on weed distribution and abundance in the experimental plots.
3. Data on population fluctuations of adult pest insects.
4. Meterological data (temperature, R.H. and rainfall).
5. Additional information on request (e.g. cost of insecticides, black plastic etc.)

Assessment of students. Students were assessed on the basis of their performance throughout the

experiment in all its phases, and on the basis of a written report on the experiment in which they were asked to develop an IPM strategy for the object agroecosystem in terms of technical, environmental and socio-economic considerations.

Results of the experiment could be broadly categorized into two types according to the treatment combinations. Either Diazinon was very much more effective than B.t. or there was little difference between Diazinon and B.t.

Faced with both these situations, students had the opportunity to express their individual attitudes toward the maximized production of the crop versus environmental protection and health.

Student responses proved to be interesting. A few did not even consider environmental issues in their discussion, choosing that strategy which reduced pest populations to the lowest level. The majority briefly mentioned environmental issues but opted for maximum economic gains from the system. Only two out of twenty students opted for environmental protection and maintained that long-term use of Diazinon was untenable.

Discussion

The above overt responses were unexpected in the light of the nature and sum total of each student's prior educational and extra-curricular experience. Three levels of response could be expected.

Level I – The cognitive/conative level. This response would arise if there had been no emphasis on biometrical analysis, and interpretation of the field experiment relied on empirical and perhaps intuitive observational judgment. The learning process would be heavily influenced by the students' cognitive style, and conclusions based on the putative experimental findings and further influenced by input from the individual's conative domain. This latter domain, in turn, could be heavily influenced by coenotropic eclecticism and possibly by some degree of pedagogism if the "correct" answer to the experiment is divined from the attitudes of the lecturer and tutor. Thus one would expect student response to be markedly coloured by environmental issues. This was demonstrably not the case.

Level II – The cognitive/scientific level. This represents the situation pertaining. The attempt to assist decision-making by emphasising the problem-solving capacity of biometrics and the computer led to some unexpected results. Firstly, an almost complete elimination of the conative style of the student. Secondly, the quasi-oracular nature of the computer created a *deus ex machina* situation inhibited exploratory motivations on the part of the students. Students were, in fact, satisfied that the computer print-outs answered all objectives in decision-making.

Level III – The holistic decision-making level. It is only at this level that the "correct" answer can be formulated. Economic analysis of treatment combinations represents but a small proportion of the total analysis required for correct decision-making. Although it was not expected that students would undertake a full-blown analysis it was expected that the factors and issues important to the system in its broadest hierarchical sence, i.e. socio-economic-political, would be discussed. This, unfortunately, was not the case.

There may be a number of reasons for this hiatus but it is possible to group the reasons into three basic categories.

1. *Lack of training in logical decision-making.* Although there is a formalised chain of decision-making links in a problem such as this, all students should have attained a level where the broad principles and issues could be discussed in qualitative terms if not in a quantitative modelling context. In particular, students originating from the Agricultural Science Faculty (almost 50% of the student number) have experienced a broad background in the rudiments of agricultural decision-making. As Smith (1973) states "... agricultural scientists are trained to deal with all the inputs of ecosystem managements, to assess outputs, and to think and act from a far wider over-view than most other professional workers". Unfortunately, Science Faculty students lack this background training.

2. *Inhibition resulting from a quasi-structured problem-solving framework.* Fensham (1972) indicated the significance that prior knowledge may have an subsequent learning in a negative sense. A corollary to this is the possibility of obtaining negative or reductionist responses from an inappropriately structured problem-solving framework. As outlined in Level II above the role of biometrics as a problem-solving tool may have grown out of all proportion to its true usefulness in the minds of the students and inhibited a more *recherche* attitude.

3. *Deficiencies in cognitive style.* Field (1971) outlined six major thinking variables or dimensions of cognitive style (conceptual differentiation, conceptual preference, originality, ideational fluency, flexibility, and category width). He found that a "science cognitive style" was associated with low flexibility and a low ideational fluency.

It seems plausible that an inability to explore the full ambit of decision-making parameters may result from a deficiency in the style or application of a student's intellectual abilities. If decision-making in its broadest and holistic sense can be considered more the role of the socio-economist than of the biological scientist, there may be some substance in the assertion that biological or agricultural science attracts students with inappropriate cognitive styles and also provides inappropriate training for decision-making.

An "ideal" discourse on decision-making would have rasied the following issues:

1. *Environmental cost.* Utilisation of conventional pest control techniques results in a cost to the environment through pollution etc. but in this case the cost accrues chiefly as a result of the elimination of biocontrol agents and the economic benefits they are capable of bestowing on the system. Experimental design allowed for this cost to be measured.

2. *Insecticide resistance cost.* Use of insecticides is to a certain extent self-defeating in that the evolution of resistance in pest populations confers a limited useful life for most compounds. The more a compound is used the shorter will be its useful life. Headley (1972) outlined the economic methods of estimating the costs of resistance and secondary pest outbreaks by capital budgeting and discounting over the entire lifetime of the strategy to obtain a nett present value criterion.

3. *Specification of the utility function.* Gains and losses of the system under different IPM strategies can be converted from monetary to utility values. Utility is a personalised assessment of preferences related to monetary gains and losses and it would be expected that each individual grower would possess a unique utility function. The role of the students is to act as a surrogate grower or grower co-operative to determine an "average" utility function suitable for the purposes of this project.

4. *Estimation of the states of nature.* As pointed out earlier in this passage conventional pest

control is based on routine, regular application of pesticide chemicals. This program is usually maintained whether required or not. For example, spraying schedules are often maintained in the absence of pest populations. Thus, the states of nature in this situation may refer to categories such as heavy, moderate, light, or no infestation.

States of nature may also refer to density levels of biocontrol agents and a host of other variables (Shoemaker, 1973). The permutations and combinations of these various states of nature will result in a formidable array of possibilities.

5. *Estimation of probability density functions.* A decision-making problem under uncertainty can be divided into three distinct classes. These correspond to: (1) no prior distribution on the states of nature; (2) a prior distribution on the states; and (3) a posterior distribution on the states. It would be a rare situation where the grower was completely ignorant of any probability distribution on the states of nature. Thus the second class outlined above would apply generally to growers with sufficient insight and experience. There are, however, econometric algorithms adequate to resolve the problems presented in class (1) and experimentation will usually produce a class (3) situation.

6. *Definition of the Choice criterion.* As Dillon (1971) states: "Logic, decision-maker's behaviour and introspection all indicate that any procedure for making a choice under conditions of uncertainty must involve two components: personal valuation of consequences and personal strength of belief about the occurrence of undertain events". Dillon advocates the use of Bernoullis' Principle which is a logical deduction from a small number of axioms which are met by a grower who wishes to be consistent and rational in his workaday decision-making.

The use of Bayes' formula (Halter and Ean, 1971) enables information about the probable states of nature in a particular situation to be incorporated into the decision framework. Final choice will depend on optimization of the physical system and maximization of the expected utility criterion (Carlson, 1970).

Conclusions

The initial execution of this course can be judged to be a successful and informative trial, although unsuccessful in terms of meeting all course objectives.

There is clearly an urgent need for quantitative analysis of three factors which appeared to have considerable effects on student achievement of course objectives 3 and 4:

1. Individual prior knowledge of the rationale underlying IPM and, in particular, of the necessity for decision-making in a socio-economic-political context.
2. Cognitive style of students undertaking the course. If it can be verified, as hypothesized, that holistic decision-making does not come easily to biological scientists, a case can be established to increase the credit point value of the course to allow for the incorporation of decision theory training as outlined above. Discrepancy in some attributes of cognitive style might be expected between students deriving from the Science and Agricultural Science Faculties. If this proves to be the norm, a bridging component in a programmed teaching strategy may be necessary to bring Science and Agricultural Science students to a comparable experiential level in this respect.
3. The potentially negative role of biometrics in the present problem-solving situation. Appropriate use of biometrics allows an independent, non-pedagogic methodology for the student to arrive at his own conclusions concerning the biophysical component of the decision-making process. The

extent to which biometric tools inhibit the full expression of the students' conative powers may be considerable.

The results of the trial underline the need for a programmed teaching approach (Billing, 1973). Assessment of program materials by provision of trials throughout the course will permit formative evaluation of the whole program, a process which would undoubtedly continue in future years.

It may also be possible, by varying this approach, to extract the full potential in decision-making from the students without appearing to unduly influence the outcomes of this decision-making.

References

ATKIN, J.M. "Research styles in science education." *Journal of Research in Science Teaching*, **5**: 69-72, 1968.

BILLING, D.E. "Programmed learning methods in advanced science courses." In *Aims, Methods and Assessment in Advanced Science Education*. Eds. D.E. Billing & D.S. Furniss, Heyden & Sons, London, 1973. Pp. 75-108.

BLOOD, P.R.B., LONGWORTH, J.W., & EVENSON, J.P. "Management of the cotton agroecosystem in Southern Queensland: A preliminary modelling framework." *Proceedings of the Ecological Society of Australia*, **9**: 230-249, 1975.

CARLSON, G.A. "A decision theoretic approach to crop disease prediction and control." *American Journal of Agricultural Economics*, **52**: 216-223, 1970.

DAVIES, P.M.C. "Towards a liberal biology." *Journal of Biological Education*, **4**: 261-278, 1970.

DILLON, J.L. "Interpreting systems simulation output for managerial decision-making." In *Systems Analysis in Agricultural Management*, Eds. J.B. Dent and J.R. Anderson, Wiley, New York, 1971. Pp. 123-151.

FENSHAM, P.J. "Prior knowledge – A source of negative factors for subsequent learning." *Research 1972*, 50-57, 1972.

FIELD, T.W. "Some dimensions of a science cognitive style and their relevance for the classroom." *Research 1971*, 19-27, 1971.

GEIER, P.W. & SPRINGETT, B.P. "The context of pest control." *Agricultural Environment*, **1**(4): 373-383, 1974.

HALTER, A.N. & DEAN, G.W. *Decisions Under Uncertainty*. South-Western Publishing Co., Cincinnati, 1971. P. 266.

HEADLEY, J.C. "Economics of agricultural pest control." *Annual Review of Entomology*, **17**: 272-286, 1972.

HILL, D.M. & SMITH, D.F. "Students' perceptions of Agricultural Science courses." *Journal of the Australian Institute of Agricultural Science*, **39**: 246-248, 1973.

NORMAN, M.J.T. "The role of the agricultural scientist in research and development." *Journal of the Australian Institute of Agricultural Science*, **39**: 51-53, 1973.

SHOEMAKER, C. "Optimization of agricultural pest management III. Results and extensions of a model." *Mathematical Biosciences*, **18**: 1-22, 1973.

SMITH, D.F. "Agriculture as a managed ecosystem." *Journal of the Australian Institute of Agricultural Science*, **39**: 247-250, 1973.

SMITH, R.F. "Integrated pest management is alive and well." Second Australian Applied Entomology Research Conference, Mildura, Working Papers. CSIRO. Agriculture and

Biological Science Branch, Victoria, 1975.

TISHER, R.P. "Trends in science education: Implications for research." *Reasearch 1971*, 5-18, 1971.

BIOLOGY AND THE KELLER PLAN

Brian G. Collins

Introduction

In recent years, an increasing number of educators has shown dissatisfaction with traditional methods of instruction which have emphasized the role of academics as disseminators of knowledge and disregarded the need to recognise the individuality of students. (Marr, *et al.*, 1960; Mertens, 1966; Gagné, 1967; Bell, 1970; Flanagan, 1971; Kormondy, 1971; Weissgerber, 1971).

As one might expect, there is no unanimity of opinion as to the best way of catering for the individuality of students. An encouraging feature which has emerged from the debate on this issue, however, has been the variety of teaching methods which have evolved. (Goldschmid & Goldschmid, 1974). In many cases, there has been a decrease in the use of lectures as a means of disseminating factual information and an increase in the use of libraries as a means of self-instruction, more use has been made of small-group discussion sessions, personalized investigative laboratory work has been introduced, and a modular approach to curriculum design has sometimes been adopted.

One method which has been introduced in recent times, in an effort to individualize student instruction and improve the efficiency of the learning process, is the Personalized System of Instruction (P.S.I.), or so-called "Keller plan". (Keller, 1966; Keller, 1967; Keller, 1968). This plan was used initially by Keller and others in association with university psychology classes, but has since been applied to education and science instruction in a variety of institutions. (Green, 1971; McFarland, 1972; Witters & Kent, 1972; Gueths, 1973; Leo, 1973; Lewis & Wolf, 1973; Nelson & Bennett, 1973; Weaver, personal communication). Keller's system has an inbuilt flexibility and potential for highly individualized instruction. It also has two features which distinguish it from most other programs involving individualized instruction. One is the requirement for unit-perfection, or mastery, by students before advancement from one unit of work to the next. The other is the use of student proctors to assist academics in the teaching program.

Despite the success that is claimed for the Keller-plan, there has been little adequate objective testing of its virtues, especially in science education, although the results which are available tend to support the intuitive feeling that student achievement is enhanced by this form of tuition. Another observation is that the data which are available tend to concentrate on comparisons between the effectiveness of conventional (lecture-laboratory) and Keller-type instruction, with few attempts being made to assess the relative values of various components of the Keller system. This dearth of information led to the development of a project, conducted in the Biology Department at the Western Australian Institute of Technology, in which the effects of mastery-learning and the use of proctors upon student achievement, retention of knowledge, and attitudes towards biology were tested for students enrolled in the animal biology section of the introductory audio-tutorial biology course. Some of the results obtained in this project are reported and discussed on the following pages.

Experimental design and administration

The portion of the project reported in this paper involved fifty students who were studying the first-year general biology unit at the Western Australian Institute of Technology. The

students had not freely elected to study biology and did not intend to proceed to advanced work in this area, but were taking the unit as a compulsory component in the first year of a Pharmacy degree course.

The experimental period began during the second half of the biology unit, and continued for a period of eight weeks. During this time, all students followed an integrated multi-experience program involving audio-tutorial tuition. Subject matter encompassed various facets of animal biology, and was broken up into a number of discrete modules or Work Units.

The main purposes of the investigations reported in this paper were to measure the effects of using (1) student instructors instead of staff instructors, and (2) a mastery-learning instructional strategy instead of a conventional audio-tutorial strategy, upon the acquisition of knowledge of biological principles, and the formation of attitudes towards (1) biology in general, (2) audio-tutorial instruction, and (3) other facets of the experimental program.

Students were assigned at random to four different treatment groups, where the tuition pattern was varied as described below, and measurements made of temporal changes in scholastic achievement and attitudes. Data obtained were then used to test the following null hypotheses:

H_0 (1) Initial knowledge of facts and principles relating to animal biology does not vary significantly between treatment groups.

H_0 (2) Initial attitudes towards biology in general, animal biology, use of student instructors, audio-tutorial instruction, frequency of testing and mastery learning do not vary significantly between treatment groups.

H_0 (3) No significant differences exist between the knowledge of facts and principles relating to animal biology by students assigned to different treatment groups at the end of the test period.

H_0 (4) No significant differences exist between the attitudes of students assigned to the various treatment groups towards biology in general, animal biology, use of student instructors, audio-tutorial instruction, frequency of testing and mastery learning at the end of the test period.

Students assigned to the first of the treatment groups continued to follow the tuition pattern that is normally adopted in the first year biology unit. Students attended large-group sessions where films were shown, talks were given by visiting experts, and matters of general interest to all students were discussed; they met with a dozen or so of their peers and one staff member in a tutorial situation; and they spent several hours per week in the audio-tutorial laboratory working with audio-tapes, study guides and demonstration materials. The second group followed the above pattern except that students were expected to demonstrate their mastery of each section of the course (usually a cluster of two or three Work Units) before proceeding to the next section. Mastery was shown by passing a Unit Test comprising twenty multiple choice items and one short-answer essay question. A level of 75% mastery was required for a pass. Three different mastery tests were available for each section of the course to cater for students who needed more than one attempt to demonstrate mastery. Testing was conducted in the time normally used for small-group tutorials by students receiving conventional instruction.

Students in the third group experienced the same kind of tuition as those in group one, except that the staff tutor was replaced by a responsible and able student enrolled in the third year of the biology degree course. The fourth group comprised students who were treated in the same way as those in group two, apart from the use of a third-year biology student instead of a staff member to administer the mastery tests.

All treatment groups were pre-tested with a one-hour achievement test, containing forty

multiple-choice items and four short-answer questions. The same test was used as a post-test after the eight week experimental period.

Of the forty multiple-choice items used in the achievement test, 40% were of the "knowledge" type and 60% "use-of-knowledge". (Bloom, 1956). The order of multiple-choice items, and the positions of correct alternatives were randomized in an attempt to prevent the establishment of set response patterns. Overall validity of the test was enhanced by ensuring that there was a fairly uniform spread of multiple-choice and short-answer items between various sections of the animal biology component of the course, and that the test items corresponded with educational objectives set for the course.

Reliability (internal consistency) of both the multiple-choice and short-answer components of the achievement test was measured by Cronbach's α (Cronbach, 1970; Mehrens & Lehmann, 1973). Reliability coefficients obtained for the post-test are reported in Table 1.

TABLE 1

Reliability coefficients for achievement post-test

Type of item	Reliability coefficient
Total multiple-choice	0.686*
Knowledge multiple-choice	0.554
Use-of-knowledge multiple-choice	0.440
Short-answer	0.709

* Coefficients calculated separately for Non-mastery and Mastery students were 0.727 and 0.535, respectively.

Indices of difficulty and discrimination were calculated for post-test multiple-choice and short-answer items (Theobald, 1964; Cronbach, 1970; Mehrens & Lehmann, 1973). Only five of the multiple-choice items failed to discriminate adequately (Ebel, 1954) and have suitable levels of difficulty; these items have been eliminated from analyses aimed at testing the validity of the hypotheses listed above.

Attitudes of students in the treatment groups towards biology in general, animal biology, audio-tutorial instruction, mastery learning, the use of student instructors, and the frequency of testing were measured by the administration of an attitude questionnaire at the times of the pre- and post-tests. Each item in the questionnaire was a statement to which the student reacted by choosing one out of four possible responses, on a summated rating scale, ranging from "strongly agree" to "agree", "disagree", and "strongly disagree" (Likert, 1932). The "feel neutral about" alternative usually provided in a Likert-type scale was not used in this questionnaire as it was felt that it provided a "soft" option for the apathetic student who did not wish to make a decision about an item. Students were instructed, on both occasions that the questionnaire was administered, to make decisions about each item if at all possible – only as a last resort were they to select "do not know".

Items were a mixture of positive and negative statements, and were worded so that agreement with an individual statement did not necessarily mean a favourable reaction (Simons, 1972). Statements were mixed randomly into the questionnaire, in order to prevent the establishment of set response patterns. Responses to the statements were scaled one to four, excluding the "do

not know" response, so that it was possible to obtain individual student attitude measures and group attitude measures towards specific items or groups of items (Veldman, 1967; Simons, 1972). Scaling was arranged so that higher measures indicated more favourable responses to questionnaire statements than did lower measures. Items that attracted a "do not know" response of >5% were excluded from subsequent analyses, thus minimising the effect of such responses upon scale means, although reliability estimates were still affected.

Cronbach's α reliability coefficients were calculated for items relating to each facet of the experimental program at the time of the post-test, as shown in Table 2.

TABLE 2

Reliability coefficients for attitude questionnaire

Scale		Cronbach's α
(a)	Biology	0.761
(b)	Animal biology	0.701
(c)	Audio-tutorial tuition	0.865
(d)	Mastery-learning*	0.933
(e)	Use of proctors*	0.893
(f)	Frequency of testing*	0.812

* α values are somewhat inflated due to the high proportion of "do not know" responses in these categories for some student groups.

The same computer program used to calculate the α values quoted above (Veldman, 1967; Docking, personal communication), provided a printout of point-biserial correlation coefficients for each questionnaire item. There was generally good correlation between each item and other items in the same item-scale.

In some cases, particularly where students had no experience of facets of the experimental program, e.g., some did not have student instructors, the percentage of "do not know" responses was high. Items in the attitude questionnaire that attracted a "do not know" response of greater than 10% were excluded from subsequent analyses.

Pre-test and post-test attitude and achievement data were analyzed by two-factor analysis of variance. Where significant differences occurred, *a posteriori* pairwise group comparisons were made using Tukey's HSD test.

Results

The results obtained are reported and statistically analyzed under the headings of the appropriate null hypotheses.

Initial knowledge of facts and principles relating to animal biology does not vary significantly between treatment groups

Achievement data for treatment groups are summarized in Table 3. These have been scaled to exclude information for the five discarded multiple-choice items. Analysis of variance indicates that there are no significant differences between groups for any of the achievement measures.

TABLE 3

Scaled pre-test achievement measures ($\bar{x} \pm$ S.D.)

	Total M/C	Knowledge M/C	Use-of-knowledge M/C	Short answer	N
Possible score	35	16	19	20	-
Pharmacy Student Groups					
01 (Staff-mastery)	8.75±4.45	4.50±2.28	4.25±2.67	0.42±0.79	12
02 (Staff-tutorial)	9.00±3.16	4.14±1.41	4.86±2.54	0.29±0.61	14
03 (Student-mastery)	9.46±3.64	4.31±2.50	5.15±1.86	0.62±0.87	13
04 (Student-tutorial)	9.91±3.88	4.00±2.24	5.91±2.12	0.36±0.67	11
Main Effects					
Staff	8.88	4.32	4.55	0.35	26
Students	9.69	4.15	5.53	0.49	24
Mastery	9.11	4.40	4.70	0.52	25
Non-mastery	9.45	4.07	5.38	0.32	25

N.B. Only approximate means are given for main effects.

TABLE 4

Pre-test attitude measures (X ± S.D.)

	Biology	Animal Biology	Audio-tutorial	Mastery	Students	Testing	N
Possible score	20	20	36	36	36	16	-
Pharmacy Student Groups							
01 (Staff-Mastery)	16.58±2.69	-	24.25±5.69	-	-	12.25±3.36	12
02 (Staff-Tutorial)	15.07±3.45	-	21.36±4.99	-	-	11.43±3.06	14
03 (Student-Mastery)	13.54±2.07	-	22.31±6.71	-	-	11.23±3.54	13
04 (Student-Tutorial)	13.00±3.44	-	24.64±6.12	-	-	11.18±2.68	11
Main Effects							
Staff	15.83	-	22.80	-	-	11.84	26
Students	13.27	-	23.47	-	-	11.21	24
Mastery	15.06	-	23.28	-	-	11.74	25
Non-mastery	14.04	-	23.00	-	-	11.31	25

N.B. Figures represent scaled data where items attracting a "non-response" of > 10% have been removed from the calculations.

Only approximate means are given for main effects.

Initial attitudes towards biology in general, animal biology, use of student instructors, audio-tutorial instruction, frequency of testing and mastery learning do not vary significantly between treatment groups.

Data relating to the attitudes of treatment groups to various facets of the experimental program, but excluding items that attracted a "do not know" response of < 5%, are summarized in Table 4. Analysis of variance indicates that the only attitude scale which produces significant treatment group differences is the one relating to "biology" ($F_{1,46} = 8.822$; $p = 0.004$).

The staff-led mastery group (group 01) had a more favourable attitude to biology than those of both student-led groups (groups 03 and 04). This is confirmed by Tukey's HSD q values: q (01 *vs.* 03) = 3.535 ($p < 0.05$); q(01 *vs.* 04) = 4.160 ($p < 0.01$).

No significant differences exist between the knowledge of facts and principles relating to animal biology by different treatment groups at the end of the test period.

Scaled post-test data is summarised in Table 5. "Mastery" students scored significantly higher than "non-mastery" when scaled scores for "total" and "knowledge" multiple choice items were used ($F_{1,46} = 6.001$, $p = 0.017$; $F_{1,46} = 9.025$, $p = 0.004$ respectively). When Tukey's analysis was applied to these data the following values were obtained: q(01 *vs.* 04) = 3.183, $p < 0.05$ for the "total" case; q(01 *vs.* 02) = 3.430, $p < 0.05$ and q(01 *vs.* 04) = 4.366, $p < 0.05$ when the knowledge items were used.

No significant differences exist between the attitudes of treatment groups towards biology in general, animal biology, use of student instructors, audio-tutorial instruction, frequency of testing and mastery learning at the end of the test period.

Scaled attitude data collected at the time of post-test are summarized in Table 6. There is a significant difference in attitude to animal biology, with the attitudes of student-led groups being generally inferior to those of staff-led groups, ($F_{1,46} = 8.618$, $p = 0.005$; q(01 *vs.* 03) = 3.0131, $p < 0.05$; q(01 *vs* 04) = 4.060, $p < 0.01$; q(02 *vs.* 04) = 2.840, $p < 0.05$).

Data for treatment groups relating to attitudes towards mastery-learning, the use of student tutors and frequency of testing were compared by t-tests. Group mean values were not significantly different for scale items relating to the use of student instructors ($t = 1.370$; d.f. = 23; $p > 0.10$), although they were for items relating to mastery learning ($t = 2.661$; d.f. = 24; $p < 0.05$) and frequency of testing ($t = 2.841$; d.f. = 24; $p < 0.01$). In both of these latter two cases, the staff-led group concerned had more positive attitudes than did the student-led group.

Discussion

Student achievement. In the past decade, numerous people have introduced "Keller-plan" programs involving mastery-learning and the use of student proctors. Several of these people have attempted to compare scholastic achievement by their students with that by students enrolled in conventional lecture-laboratory classes which followed the same curriculum. Evidence that has been obtained – much of it extremely subjective and lacking a comprehensive statistical basis – supports the view that achievement by students in "Keller-plan" classes is at least equivalent (Lewis & Wolf, 1973), and often superior (Keller, 1968; McMichael & Corey, 1969; Alba & Pennypacker, 1972; Born, *et al.*, 1972; Morris & Kimbrell, 1972; Witters & Kent, 1972), to that by students experiencing tradit-

TABLE 5

Scaled post-test achievement measures ($\bar{x} \pm$ S.D.)

	Total M/C	Knowledge M/C	Use-of-knowledge M/C	Short answer	N
Possible score	35	16	19	20	-
Pharmacy Student Groups					
01 (Staff-mastery)	22.25±4.63	12.00±1.95	10.25±3.08	8.75±3.93	12
02 (Staff-tutorial)	19.21±3.81	9.79±2.12	9.43±2.50	8.14±3.46	14
03 (Student-mastery)	21.69±3.71	10.85±2.15	10.85±2.16	8.46±2.82	13
04 (Student-tutorial)	18.73±5.16	9.18±2.86	9.55±3.01	8.00±4.63	11
Main Effects					
Staff	20.73	10.89	9.84	8.45	26
Students	20.21	10.01	10.20	8.23	24
Mastery	21.97	11.42	10.55	8.61	25
Non-mastery	18.97	9.48	9.49	8.07	25

N.B. Only approximate means are given for main effects.

TABLE 6

Post-test attitude measures (X ± S.D.)

	Biology	Animal Biology	Audio-Tutorial	Mastery	Students	Testing	N
Possible score	20 (24)	20	36	36	24	16 (24)	-
Pharmacy Student Groups							
01 (Staff-Mastery)	16.50±2.47 (19.17±2.48)	17.00±3.44	25.92±7.06	27.92±5.12	-	14.08±1.68 (20.92±2.02)	12
02 (Staff-Tutorial)	14.14±3.78 (16.43±4.29)	16.07±1.69	21.00±7.16	-	-	-	14
03 (Student-Mastery)	14.62±3.66 (16.38±4.07)	14.69±3.04	20.69±3.04	22.85±4.43	16.77±3.52	11.62±2.50 (17.08±4.30)	13
04 (Student-Tutorial)	15.09±2.30	13.91±2.30	22.18±5.20	-	14.27±5.29	-	11
Main Effects							
Staff	15.32 (17.80)	16.54	23.46	27.92	-	14.08 (20.92)	26
Students	14.85 (16.51)	14.30	21.44	22.85	15.52	11.62 (17.08)	24
Mastery	15.56 (17.78)	15.85	23.30	25.39	16.77	12.85 (19.00)	25
Non-Mastery	14.62 (16.54)	14.99	21.59	-	14.27	-	25

N.B. Figures represent scaled data where items attracting a "non-response" of > 10% have been removed from the calculations. Values in parentheses incorporate data for additional questionnaire items ("Biology" – one item; "Testing" – two items) which were valid on the post-test questionnaire, but not on the pre-test questionnaire.
Only approximate means are given for main effects.

ional lecture-laboratory instruction.

Investigations which have been conducted have usually concentrated on assessing the virtues of the "Keller-plan" *in toto*, rather than testing the relative merits of its many components. Little, if any, attention has been directed at evaluating whether it is the use of students instead of staff tutors that is the most important factor in the "plan" . . . or whether it is the mastery-learning strategy itself, or, perhaps, other factors. In addition, there seems to have been no attempt made to measure the effectiveness of various "Keller-plan" features when incorporated into a teaching strategy involving audio-tutorial instruction. This study was therefore developed in an attempt to provide an expanded analysis of the "Keller-plan", looking in particular at some of the above aspects.

Treatment groups investigated in the experimental program were composed of randomly assigned students and were statistically equivalent on the basis of all achievement measures used at the time of the pre-test (Table 3). These groups were found to have acquired significant quantities of knowledge and developed their ability to apply this knowledge by the end of the experimental period of eight weeks. When simple comparisons of treatment groups are made, it can be seen that mean achievement scores for groups experiencing mastery-learning were superior to those for other groups on all categories of test items. In other words, the mastery-learning strategy seems to be superior to the non-mastery strategy, regardless of whether a student or staff tutor is used. Statistical comparison of group scores, however, indicates that not all groups had progressed to the same degree, except in the case of performance on "short-answer" and "use-of-knowledge" multiple-choice items. Group 01, the treatment group with a staff tutor which followed a mastery-learning strategy, performed significantly better on "total" multiple-choice items than did group 04, the group with a student tutor that followed a conventional audio-tutorial strategy. This difference was primarily due to superior performance on "knowledge" multiple-choice questions. A significant difference was also found to occur between the mean "knowledge" scores for group 01 and group 02, the group with a staff tutor which received conventional audio-tutorial instruction, although this difference was not reflected in the "total" multiple-choice scores.

This evidence lends support to the view that mastery-learning, *per se*, has a positive effect upon student achievement. A fairly clear indication has also been given that the instructional strategy involving student tutors and non-mastery learning is the least effective of those reported in this paper, although the student-mastery approach is not inferior to the two methods involving staff instructors. Various reasons can be advanced for these observations. The most obvious, perhaps, is that the frequent learning and testing associated with unit-mastery tests leads to a greater retention of factual information by mastery groups that is reflected in enhanced performance on "knowledge" items at the time of the post-test. The lack of significant treatment group differences on "use-of-knowledge" and "short-answer" items is a little surprising, however, in view of the limited, but fairly convincing, results obtained by other workers. Sheppard & MacDermott (1970), for instance, have demonstrated superior achievement by mastery students on both multiple-choice and essay (short-answer) test items. Born, *et al.* (1972) also showed that student performance on fill-in and short-answer items was enhanced by the "Keller-plan" strategy. Morris & Kimbrell (1972) obtained superior performance scores on "use-of-knowledge" type questions from their mastery-learning students, although this did not occur for "knowledge" items.

It must be remembered that the results obtained by all three groups of workers mentioned above were obtained when comparing mastery groups with control groups receiving lecture-laboratory instruction where there was a minimum of student-instructor interaction. On the other hand, results reported in this paper were obtained from students who all followed the same basic audio-

tutorial instructional pattern where there was always a considerable amount of student-instructor contact, and an opportunity for self-pacing. Perhaps this led to a fairly uniform degree of understanding of the subject matter which was reflected in similar group performances on "short-answer" and "use-of-knowledge" items which were designed to test understanding rather than simple recall.

Student attitudes. Initial attitudes of students to most facets of the experimental program did not differ significantly between the four test groups. In fact, the only detectable differences were those relating to the "biology" category, where group 01 (staff-mastery) had more positive attitudes than did groups 03 (student-mastery) and 04 (student-tutorial).

The above differences had disappeared by the end of the test period, suggesting that the use of student-instructors may have had a positive effect on the attitudes of first-year students towards biology. This trend is not repeated, however, when attitudes to other facets of the program are considered. In fact, fairly conclusive evidence has been presented to indicate that groups with student-instructors develop attitudes towards animal biology, mastery-learning and frequency of testing that are less positive than those with staff-instructors. Data available give little indication that students experiencing mastery-learning develop any different attitudes to biology, animal biology and audio-tutorial instruction than do those attending conventional tutorials. In this regard, the reported findings of other workers (Born, *et al.*, 1972; Morris & Kimbrell, 1972; Witters & Kent, 1972), where "Keller-plan" learning was found to promote the development of more positive attitudes, are not supported. It should be stressed again, however, that the findings of the workers mentioned above were reached after drawing comparisons between "Keller-plan" and conventional lecture-laboratory classes, whereas the results reported in this paper have been gained with students who all experienced the same basic personalised, audio-tutorial method of instruction.

References

ALBA, E., & PENNYPACKER, H.S. "A multiple change score comparison of traditional and behavioural college teaching procedures." *Journal of Applied Behaviour Analysis*, **5**: 121-124, 1972.

BELL, P.E. "An exploration of alternative methods for teaching large sections of general education biology." *Journal of Research in Science Teaching*, **7**: 363-369, 1970.

BLOOM, B.S. *Taxonomy of Education Objectives: Handbook I, Cognitive Domain.* David McKay Company, New York, 1956.

BORN, D.G., GLEDHILL, S.M., & DAVIS, M.L. "Examination performances in lecture-discussion and personalized instruction courses." *Journal of Applied Behavior Analysis*, **5**: 33-43, 1972.

CRONBACH, L.J. *Essentials of Psychological Testing.* Harper & Row, New York, 1970.

EBEL, R.L. "Procedures for the analysis of classroom tests." *Educational and Psychological Measurement*, **14**: 352-364, 1954.

FLANAGAN, J.C. "Individualizing Education." In Weissgerber, R.A. (Ed.), *Perspectives in Individualized Learning.* Peacock, Illinois, 1971.

GAGNÉ, R.M. "Instruction and the conditions of learning." In *Instruction: Some Contemporary Viewpoints,* (Ed. L. Siegel). Chandler, San Francisco, 1967.

GOLDSCHMID, B., & GOLDSCHMID, M.C. "Individualized instruction in higher education: A review." *Higher Education*, **3**: 1-24, 1974.

GREEN, B.A. "Physics teaching by the Keller plan at Massachusetts Institute of Technology." *American Journal of Physics*, **39**: 764-775, 1971.

GUETHS, J.E. "A proctor's course for the Keller plan." *Journal of College Science Teaching*, **2**: 30-32, 1973.

KELLER, F.S. "A personal course in psychology." In R. Ulrich, T. Stachnik, & J. Mabry (Eds.) *The Control of Behavior*. Scott, Foresman, Glenview, Illinois, 1966.

KELLER, F.S. "Neglected rewards in the educational process." Proceedings of the 23rd American Conference of Academic Deans, Los Angeles, 1967.

KELLER, F.S. "Goodbye teacher." *Journal of Applied Behavior Analysis*, **1**: 78-89, 1968.

KORMONDY, E.J. Preface to J.G. Creager & D.L. Murray (Eds.) *The use of modules in college biology teaching*. C.U.E.B.S., 1971.

LEO, MICAH WEI-MING. "Chemistry teaching by the Keller plan." *Journal of Chemical Education*, **50**: 49-50, 1973.

LEWIS, D.K., & WOLF, W.A. "Implementation of self-paced learning (Keller method) in a first year course." *Journal of Chemical Education*, **50**: 51-56, 1973.

LIKERT, R. "A technique for the measurement of attitude." *Archives of Psychology*, **22**: 140-143, 1932.

McFARLAND, S.J. "Effects of a program of self-paced proctored instruction on learning and attitudes at the university level." *Dissertation Abstracts*, **32**(*10-A*): 5615, 1972.

McMICHAEL, J.S., & COREY, J.R. "Contingency management in an introductory psychology course produces better learning." *Journal of Applied Behavior Analysis*, **2**: 79-83, 1969.

MARR, J.N. "The contribution of the lecture to college teaching." *Journal of Educational Psychology*, **51**: 277-284, 1960.

MEHRENS, W.A. & LEHMAN, I.J. *Measurement and Evaluation in Education and Psychology*. Holt, Rinehart & Winston, New York, 1973.

MERTENS, T.R. "Programmed instruction and the changing role of the faculty member." *American Biology Teacher*, **27**: 779-782, 1966.

MORRIS, C.J., & KIMBRELL, G. McA. "Performance and attitudinal effects of the Keller method in an introductory psychology course." *Psychological Record*, **22**: 523-30, 1972.

NELSON, T.F., & BENNETT, M.L. "Unit size and progress rates in self-pased instruction." *Journal of College Science Teaching*, **3**: 130-133, 1973.

SHEPPARD, W.C. & MacDERMOT, H.G. "Design and evaluation of a programmed course in introductory psychology." *Journal of Applied Behavior Analysis*, **3**: 5-11, 1970.

SIMONS, L. "A comparison of the relative effectiveness of written scripts and audio tapes in teaching college biology." Ph.D. Dissertation, Syracuse University, 1972.

THEOBALD, J.H. *An Introduction to the Principles of Classroom Testing*. Hall, Melbourne, 1964.

VELDMAN, D.J. *Fortran Programming for the Behavioural Sciences*. Holt, Rinehart & Winston, New York, 1967.

WEISSGERBER, R.A. *Perspectives in Individualized Learning*. Peacock, Illinois, 1971.

WITTERS, D.R., & KENT, G.S. "Teaching without lecturing: evidence in the case of individualized instruction." *Psychological Record*, **22**: 169-175, 1972.

SCIENTIFIC METHODOLOGY : A FEYERABENDIAN VIEW

Barry C. Newman

Introduction

For one reason or another, science teacher educationists have assumed that science education is concerned with the nature of science, one important feature of the nature of science being scientific methodology (Collette, 1973, pp. 1–24; Renner and Stafford, 1972, pp. 1–36; Sund and Trowbridge, 1973, pp. 1–24; Tisher, Power and Endean, 1972, pp. 23,24), to cite just a few. What constitutes scientific methodology, however, has sometimes been a point of heated debate (Lakatos and Musgrave, 1970). Inductivism, Falsificationism and versions of each have been proposed only to come under fire.

While there has been some recognition by science teacher educationists of the contributions to the debate of some philosophers of science, Popper and Kuhn, for example, little attention has been given to the works of Imre Lakatos or Paul K. Feyerabend. The purpose of this paper is to remedy in part this situation by focusing on Feyerabend. But first a word on Popper, Kuhn and Lakatos.

Popper's 'Falsificationism'

For some tears now, Popperian 'falsificationism' has been propounded as at least a significant ingredient of scientific methodology, to be recognised as such by science teachers and to be passed on by them to pupils of science (for example, Jevons, 1969, p. 32; Martin, 1972, pp. 32–40, 42, 43; Phillips, 1968; Tricker, 1967, pp. 22, 23). It is probably the case that many more educationists have implicitly appealed to Popper's methodology than have explicitly referred to his name.

Popper's position, detailed in *The Logic of Scientific Discovery* (1972a) [first edition, appearing in German in 1934], in outline is as follows: What one discovers in science is that which is false, rather than that which is true. The logic of scientific discovery is (i) the formulation of a universal generalisation (a theory or an hypothesis); (ii) the logical deduction of singular statements (or hypotheses) that would be logically inconsistent with the universal generalisation, and (iii) the testing of these singular statements and hence the theory by deciding to accept relevant basic statements which assert that certain observable events occur. If a basic statement that one accepts is equivalent to a singular statement deduced as inconsistent with the theory, then the theory is falsified. If a theory has not been falsified, but none the less tested, it is 'corroborated' to the extent to which it has been tested. Universal generalisations are never verified, they can only be falsified or corroborated. In order to meet certain objections to the practice of this methodology (e.g. the objection that in principle it is always possible to challenge a basic statement – basic statements are themselves theory-laden) Popper proposes the use of certain procedural rules. These methodological rules enable the falsifying process to work.

Popper does not believe that his methodology necessarily describes how scientists act; rather, it prescribes how they ought to act. It is seen by him as a logical challenge to any form of inductivism as a methodology for arriving at the truth. For a recent attempt at a fortification and a modification of Popper's view, see Maxwell (1972).

Kuhn's 'scientific revolutions'

A more recent portrayal of scientific methodology – Kuhn's *The Structure of Scientific Revolutions* (1970a) [first edition 1962], has also gained the attention of a number of science teacher educationists (Christian, 1973; Hurd, 1969, p. 24; Macklin, 1973; Young, 1974 etc.). It is likely that an increasing appeal will be made to Kuhn's theories, not only by those interested in science teacher education, but by those interested in the developmental aspects of any enterprise.

Unlike Popper's account, Kuhn's purports to be both prescriptive and descriptive, and whereas Popper highlights the falsification of theories, Kuhn surveys the general notion of 'scientific progress'. Rather than simply producing a methodology for research, Kuhn points to salient features in the history of science, the main ones being 'the scientific community', 'paradigms', 'commitment to paradigms' or 'normal science', 'anomalies', 'crisis', 'paradigms being questioned' or 'extraordinary science', 'revolution' and the alogical 'Gestalt switch'. In a mature science, most of the time, most of the community are engaged in the activities of 'normal science'. The members of the scientific community do not question anything of the disciplinary matrix to which they are committed but work within a single paradigm, solving puzzles of that paradigm. As 'normal science' progresses, however, anomalies begin to appear, several competing theories may emerge, and different commitments to the paradigm are made: a time of extraordinary science has arrived. If the crisis is severe enough revolution is likely to follow. Competing paradigms will emerge that will not only be incompatible, but also often incommensurable. In the end, it is very likely that a new paradigm will head the field, a 'Gestalt switch' will have been made, and the time for normal science will have returned. Progress has been made in moving from one paradigm to another because there has been a move from a worse to a better paradigm as viewed by the scientific community. The community will not say there has been anything but progress. Progress is also made as 'puzzles' are solved during times of normal science.

For recent critiques of Kuhn's position, see Berkson (1974) and Novakovic (1974). Both comment, in addition, on Popper, Lakatos and Feyerabend.

Lakatos' 'sophisticated methodological falsificationism

Compared to Kuhn or Popper, Lakatos has received only slight attention from science teacher educationists. For example, he is mentioned only in a note and in a bibliography in Martin's *Concepts of Science Education* (1972, pp. 162, 170) and in a paragraph in Christian's 'Verification and Falsification in Science' (1973) but only in terms of an explication of a Popperian view. Tawney refers to him, but only in comparing Popper and Kuhn (1974, p. 25). Christian's reference to Popper is a little odd since it is really only Lakatos that he quotes from in this connection. It is true that in his 'Falsification and the Methodology of Scientific Research Programmes' (1970) Lakatos appeals to some of Popper's ideas on theory preference (see Popper, 1972a, pp. 50, 419; 1972b, pp. 13–16, 21, 22) but in this work he makes a significant contribution of his own and is deserving of more attention.

Although the methodology proposed by Lakatos is fundamentally prescriptive, he also believes that it is supported by historical evidence, though one has to rationally reconstruct some of the evidence in order to see this. For Lakatos there is no falsification of theory 'T' before the emergence of a better theory 'T,'. 'T,' has empirical content over 'T', explains the previous success of 'T' and some of its excess content over 'T' is corroborated. With this new notion of 'falsification' Lakatos describes series of theories as theoretically progressive, empirically progress-

ive, progressive or degenerating. If each new theory in a series has excess empirical content over its predecessor, the series is theoretically progressive. If some of this excess is also corroborated, the series is empirically progressive. If a series of theories is both theoretically and empirically progressive the series of theories is progressive; if it is not, then it is degenerating. Scientific Research Programmes are concerned with series of theories (problem-shifts). In any such program there will be a number of methodological rules. One rule will be concerned with the 'hard core' of the program: the hard core is to be regarded as 'irrefutable'. Other rules will be concerned with the 'protective belt' of the program: the theories of the protective belt may change in particular ways. Other rules will relate to the assessment of evidence, etc.

As regards so-called 'crucial experiments', Lakatos believes that they are generally misunderstood. Crucial counter-evidence can be recognised as such among the score of anomalies only *with hindsight* in the light of some superseding theory.

For an amplification of Lakatos' views on crucial experiments see Lakatos (1974). For a recent critique of Lakatos' view see Quinn (1972).

Feyerabend's amethodology

Feyerabend shares Lakatos' fate; he is practically unknown. He is mentioned by Christian (1973) and by Macklin (1973) but only incidentally. Martin refers to him but only in his notes and bibliography (1972, pp. 165, 166, 169). It is not as though he has only written a little on scientific methodology: see Feyerabend (e.g. 1962; 1965; 1968; 1970a; 1970b; 1970c). Nor is it that he hasn't been engaged in public debate: for example with Popper, Kuhn, Lakatos and others (Feyerabend 1970b), and with Lakatos (Feyerabend, 1973). It may be that he is difficult to read or understand, or that his position is without even plausible foundation, or that what he writes and says is 'distasteful'.[1]

(i) *Feyerabend on Popper.* There is little of similarity between Popper's and Feyerabend's views and Feyerabend spends little time comparing them. He is opposed to the notion that theories can be so quickly judged and then possibly dismissed as easily as Popper's methodology suggests (1970b, p. 215). He is profoundly unhappy with Popper's assertions of the rational character of science that is claimed to exist not only in the falsification of theories, but also in their comparison (1970b, p. 218).

(ii) *Feyerabend on Lakatos.* Feyerabend believes that one of the significant contributions to the debate is Lakatos' claim that in any period of scientific endeavour, theories are proliferated (as in a series of theories in a research program) while at the same time, research programs are protected (degenerating research programs must be allowed to remain for some time, they might turn into progressive ones) (1970b, p. 203 note 2, p. 211). The idea that proliferation and protection occurs at the same time is used by Feyerabend in an attack upon Kuhn (1970b, pp. 211–214).

However, Feyerabend asserts that as a methodology, Lakatos' prescription is fundamentally faulty. Since degenerating research programs might turn into progressive ones at *any* time, one never knows when one should give them up (1970b, pp. 214–216). Furthermore, there is always a plurality of theories and therefore Lakatos should give up his distinction between research programs and individual theories and should widen his analysis to take account of incommensurable theories (1970b, p. 203 note 2, p. 211).

(iii) *Feyerabend and Kuhn.* Feyerabend and Kuhn agree on a number of issues. Firstly, both

believe that the overthrow of one theory requires at least one other theory (Feyerabend, 1962, p. 32; 1970b, pp. 207–211; Kuhn, 1970a, pp. 70–75). Secondly, both believe that there are some theories which are incommensurable (Feyerabend, 1970b, pp. 219–222; Kuhn, 1970a, p. 103). Thirdly, both agree that, as a partial consequence, there are some significant aspects of scientific progress which are irrational (Feyerabend, 1970b, pp. 214, 215) or at best alogical (Kuhn, 1970b, pp. 259–266), and fourthly, both deny the charge of relativism and agree that theory choice is often a rational choice (Feyerabend, 1970b, pp. 228, 229; Kuhn, 1970a, pp. 205–207).

For all that agreement, Feyerabend contends that he is in basic disagreement with Kuhn (1970b, p. 197), that Kuhn's descriptive account of Single Paradigm – Multi-Paradigm – Single Paradigm is historically false (1970b, pp. 207, 208, 211–214). It is his model and not Kuhn's that is in better agreement with the history of science (1970b, p. 210). Historically, Feyerabend argues, proliferation of theories does not begin when a revolution begins, it precedes it, and is necessary in order for a revolution to occur. Not only does proliferation of theories precede revolution, it is there all the time. Science is not a temporal succession of normal periods and periods of proliferation of theories, it is their juxtaposition. These facets of scientific activity do not occur as *periods*, but as *components*. It is the *interaction* of normal science with a philosophical part (paradigm questioning, proliferation of theories etc.) that advances it. It is only because the normal component is large, well entrenched, and relatively resistant to change, that revolutions are so noticeable. The new material thrown up by the philosophical component is publicised, by those who have an interest in it, in opposition to those entrenched views held by the conservatives – the party of the normal component. In the end, the normal elements may change, but if they do, it will be for any number of arational reasons (a leading conservative dies, a younger generation cannot be bothered to follow its elders etc.). A revolution has then occurred, a change in the normal component has been brought about – revolution being the outward manifestation of these changes (1970b, pp. 211–214).

(iv) *Feyerabend's normative stance*. In spite of his concern with Kuhn's so-called historical account and his replacement of this account with his own, Feyerabend's 'methodology' for scientific activity is fundamentally a prescriptive one and not descriptive. At the same time, however, he is not beyond appealing to history to support his position or to illustrate it.

His model is purportedly a product of his desire for scientific activity to fulfil two aims. As opposed to other epistemologies, an anarchistic epistemology (and hence an anarchistic philosophy of science) "is not only a better means for *improving knowledge* or of understanding history, it is also more appropriate for a free man to use" (1970a, p. 21, my emphasis). Knowledge is best improved, if one aims at maximum *testability* of knowledge (1965, pp. 223–225). "The interests of man, and, above all . . . his freedom (freedom from hunger, despair, from the tyranny of constipated systems of thought)" (1970a, p. 76) are best served if we reform "the science of critical rationalism" and make it "more anarchistic and more subjective" (1970a, p. 76). Understandably, Feyerabend does not support the normal component of scientific activity, but rates it as anti-humanitarian (1970b, pp. 210, 212).

(v) *The principle: 'anything goes'*. Feyerabend's 'methodology' – his model for scientific activity – is no methodology at all; or, to put it another way: *any method and every method will do*. Hence the title to one of his works – 'Against Method' (1970a).

Feyerabend argues that "there is not a single rule, however plausible and however firmly grounded in epistemology, that is not violated at some time or another" and "we see that (such

violations) are necessary for progress" (1970a, p. 22).

> Considering any rule, however "fundamental", there are always circumstances when it is advisable not only to ignore the rule, but to adopt its opposite. For example, there are circumstances when it is advisable to introduce, elaborate, and defend ad hoc hypotheses, or hypotheses which contradict well-established and generally accepted experimental results, or hypotheses whose content is smaller than the content of the existing and empirically adequate alternatives, or self inconsistent hypotheses, and so on. There are even circumstances . . . when *argument* loses its forward-looking aspect and becomes a hindrance to progress (1970a, p. 22).

"The idea of a fixed method . . . arises from too naive a view of man and his social surroundings . . . there is only one principle that can be defended under all circumstances, and in *all* stages of human development. It is the principle: anything goes" (1970a, pp. 25, 26).

Feyerabend suggests two other principles consistent with the principle 'anything goes'; the principle of proliferation: "invent and elaborate theories which are inconsistent with the accepted point of view, even if the latter should happen to be highly confirmed and generally accepted" (1970a, p. 26) and the principle of tenacity: "select from a number of theories the one that promises to lead to the most fruitful results, and . . . stick to this one theory even if the actual difficulties it encounters are considerable" (1970b, p. 203). With respect to the latter principle, the emphasis seems to be on the survival of a theory rather than its selection.

Presumably, these two principles are consistent with the 'anything goes' principle, in that (a) the principle of proliferation aims at not allowing 'established' theories or facts to remain unchallenged, and (b) the principle of tenacity aims at allowing any 'successfully' attacked theory to recover. In spite of the principle of tenacity, however, or as Feyerabend argues, because of it (1970b, p. 205) theories can be removed (presumably, they can also be revived). They can not be removed because they are 'successfully' attacked by facts – the principle of tenacity forbids that; they can only be removed, if, in response to the principle of proliferation, alternative theories are provided which accentuate the difficulties of the theories to be removed while promising means for their solution.

Now Feyerabend, while often appealing to the principle of proliferation, rarely mentions the principle of tenacity. It may be that discussion of the former principle better suits his polemical purposes, or it may be that he considers the principle of proliferation to entail the principle of tenacity (see 1965, p. 224, where Feyerabend appears to subsume the principle of tenacity under the principle of proliferation). Proliferation of theories may be taken to imply that older theories cannot be permanently eliminated.

As an essential part of any *critical* empiricism, and as an essential part of a humanitarian outlook, Feyerabend suggests that proliferation of theories is a necessary principle. From a practical point of view, relevant evidence for the test of a theory can often be unearthed only with the introduction of an incompatible alternative theory (1970a, p. 26). From a humanitarian point of view, "proliferation means that there is no need to suppress even the most outlandish product of the human brain. *Everyone may follow his inclinations* . . . The interplay between proliferation and tenacity also amounts to the continuation, on a new level, of the biological development of the species" (1970b, p. 210).

(vi) *'Theories' versus 'facts'.* Now, as indicated above, the principle 'anything goes' implies that

theories should be invented such that they are not only inconsistent with other theories, but also such that they are inconsistent with 'experiments', 'facts' and 'observations'. The argument goes: no single theory ever agrees with all the known facts in its domains, so there shouldn't be any concern about introducing other theories with this 'problem' (1970a, p. 36). There are quantitative disagreements (e.g. the special theory of relativity, and Miller's evidence, in the 1920s, for the existence of an ether) (1970a, p. 37 note 10) and there are qualitative disagreements (e.g. electrodynamical theory contains the absurd consequence that the motion of a free particle is self-accelerated) (1970a, p. 39). Sometimes these disagreements are dealt with by being ignored (1970a, p. 39), and sometimes, *ad hoc* approximations are made (1970a, p. 41). If every theory is 'in trouble', why not introduce another theory, whatever its relationship with the 'facts', which in the course of time might turn out to be less 'in trouble'.

Furthermore, it may not be that the theory is 'incorrect'; it could be that the 'evidence', 'facts' or 'observations' are contaminated. Difficulties could lie in the 'observation terms', 'the sensory concepts', 'the auxiliary sciences' or 'the background speculation' and that could be why the theory is inconsistent with the evidence (1970a, p. 44).

Adopting an Hegelian world view, Feyerabend argues for the 'motion' of concepts, the 'motion' of objects, and the 'motion' of concept: object systems, as necessarily being part of the critical empiricism which proposes confrontations of theories with facts. "Do not work with stable concepts . . . Do not be seduced into thinking that you have at last found the correct description of 'the facts'. We must . . . proceed dialectically . . . by an *interaction* of concept and fact (observation, experiment, basic statement, etc.)", (1970a, p. 36). Be prepared to let 'the fact' and 'the concept' change. 'A stone falling in an absolute reference frame' may become 'a stone falling in a reference frame relative to an observer's reference frame' or *vice versa*, or both may become something else. A 'continuous energy system' may become a 'quantized energy system', and electrons may become no longer 'objects of definite mass, size and locality', but 'wave functions' or *vice versa*, or both may become a new 'fact' with a new 'concept'. Particularly is it important to criticise commonly used concepts. Being commonly used they are firmly believed to be 'adequate', 'correct', 'proper' and 'true', and are rarely criticised in practice. One must invent "a new conceptual system, a new theory that clashes with the most carefully established observational results and confounds the most plausible theoretical principles" (1970a, p. 45). Science cannot exist without such moves. They are legitimate and much needed (1970a, p. 45).

In order to highlight some of the above issues, Feyerabend makes a lengthy analysis of Galileo's tower argument (1970a, pp. 45–70). As a proposed refutation of Copernicus' view of the rotation of the earth, an argument is introduced pointing out that when a stone is allowed to fall from a tower, it falls straight down, and does *not* land some distance 'behind' the tower, the latter possibility being deemed necessary if the earth was moving. Feyerabend demonstrates how Galileo 'defuses' this argument; defuses rather than refutes, because a changing conceptual system is involved and not a logical refutation of a conceptual system. Galileo examines the argument to discover those 'natural interpretations' (those mental operations which follow so closely upon the senses, that a separation is difficult to achieve) which are responsible for the 'contradiction'. He then *replaces* these natural interpretations (involving absolute motion) with others (involving relative motion) and so replaces an old observational language with a new one. To do this he uses *propaganda*. He introduces an analogy involving 'relative motion' in the eyes of most observers, and persuades the objector to adopt a relative motion viewpoint at least as regards observers on a rotating earth. Feyerabend suggests that Galileo used *ad hoc* hypotheses in that as he gradually accepted the Copernican view, he endeavoured to avoid difficulties with the tower argument by

simply adopting a circular law of inertia and the idea of relativity of motion. He had to be *ad hoc* one way or another, either with an old theory or a new theory. "Hence, as one cannot help being *ad hoc*, it is better to be *ad hoc* with respect to a new theory, for a new theory, like all new things, will give a feeling of freedom, excitment, and progress" (1970a, p. 69) – Feyerabend's 'humanitarian' outlook again.

Feyerabend answers the challenge that he is confusing observation with theory on the one hand, and a context of discovery with a context of justification on the other hand, by pointing out that such sharp distinctions do not exist, or if they do, they reflect a temporary stasis of the process of research, and the situation involving such distinctions should be eliminated as quickly as possible (1970a, pp. 70–72). It is true that theories are often tested by experience, but experience is also analysed with the help of theories. Experience as well as theories contains 'natural interpretations' which are abstract and contain metaphysical ideas such as the idea of an observer – independent existence. Observations are not discovered, although causally they are related to the empirical world. "An observation sentence is distinguished from other sentences of a theory, not . . . by its content, but by the cause of its production or by the fact that its production conforms to certain *behavioural* patterns" (1962, p. 36). It is never the case that observation statements are certain and theories only in jeopardy. "Research at its best is *interaction* between newer theories which are stated in an explicit manner and older views which have crept into the observational language" (1970a, p. 70).

(vii) *Incommensurable theories*. Feyerabend recognises that opponents may claim that in spite of the difficulties in scientific procedure, in spite of the actual history of science, and regardless of whether or not it might be constrictive of man's freedom and humanity, a scientific methodology, say similar to Popper's or Lakatos', might still be more 'successful' than Feyerabend's 'amethodology', and a reconstructed (but not too reconstructed) scientific enterprise more rational than Feyerabend thinks possible. He answers such opponents by pointing to the incommensurability of certain theories (1970a, p. 81–90).

He uses the theory of special relativity (SR) and classical mechanics (CM) as examples. In CM, the terms 'mass', 'length' and 'time' are associated with values that are *independent* of the velocity of an object relative to an observer. In SR, the terms 'mass', 'length' and 'time' are associated with values that are *dependent* on the velocity of an object relative to an observer. In CM all events are embedded in three-dimensional Euclidean Space, with an added time co-ordinate. In SR, all events are embedded in a four-dimensional Riemannian continuum. In addition, the terminology for both CM and SR is comprehensive. It pervades the theories. Since the comprehensive terminology of CM and SR is different for CM and SR, no statements of the theories can be shared. It is impossible therefore, for one theory, say SR, in replacing CM, to meet such formal requirements as: it has to repeat the successful consequences of the older theory, it has to deny its false consequences, and it has to make additional predictions. These formal requirements cannot be met in the sense that the theories are radically different theories. The consequences and predictions for one theory are not the same type of consequences and predictions for the other theory. They cannot be compared. "The consequence classes of SR and CM are not related in any way. A comparison of content and a judgement of verisimilitude cannot be made" (1970a, p. 84). It is not that their domains are different, one theory *replaces* the other, such that it is a *world* view replacing another, different *world* view. The formal requirements of both Popper and Lakatos (as they concern theories within research programs) cannot be met, in the case of incommensurable theories, because both require that a theory replace another theory, having comparable content

(Popper, 1972b, p. 14; Lakatos, 1970, p. 116).

It is not the case that one set of terms within CM has meanings more 'natural' than the meanings of the other set of terms within SR. One language may prove to more difficult than another, but both may be as 'unnatural' or as 'natural' as each other. Nor is translation possible. One can show where one meaning replaces (occupies the domain occupied by) the other, but one is not entitled to say one equals the other, or stands for the other. To say that 'mass' in SR equals 'mass' in CM, if the velocity of the observer is zero, is only to say that the values are the same, but not the meanings.

All that one can ask of theories, says Feyerabend, is that they should give us a correct account of the world as seen through their *own* concepts, and that *in the final analysis,* they be seen to be self-consistent. There is a certain sense therefore, in which incommensurable theories can be compared, and sometimes a choice made between them, and this entails the notion of the crucial experiment. The *one* experiment will be viewed differently by SR and CM. Both will give rise to observation statements, different from one another, because different meanings are involved, even though the same terms might be used and even though these statements are derived from the one experiment. The observation statement of one may be found to be inconsistent with the theory within which it was formed, but the other observation statement may be consistent with the theory within which it was formed. On this basis, one could eliminate one theory but not the other. Such a 'refutation', involving *incommensurable* theories is a weak refutation (1970a, p. 90).

One can understand why incommensurable theories are significant for Feyerabend, not only because they reveal the impossibility of applying certain methodological rules, but also because they are the sorts of theories that one would expect to be developed as a result of 'concept': 'fact' mobility. If one wants to question the facts, and question the concepts, then in the proliferation of theories incommensurable theories are very likely to result. In addition, however, Feyerabend makes an appeal to incommensurable theories in an attack upon formal accounts of explanation and reduction (1962, pp. 28–97). Such accounts, by Hempel and Oppenheim (1948) and Nagel (1949), for example, require meaning invariance for the same terms and strict logical connections. However, if commensurable theories are involved, meaning invariance is not possible and, consequently, strict logical deductions are likewise not possible.

(viii) *Irrationality in theory choice.* Whether incommensurable theories are involved or not, however, the principle 'anything goes' always holds. There are no constricting methodological rules which are the final arbiters in matters of theory propagation or theory choice. In the final analysis (although it is particularly true in the case of incommensurable theories) "what remains are aesthetic judgements, judgements of taste, and our own subjective wishes" (1970a, p. 90). This is not to claim, however, that such matters are completely beyond the reach of argument (1970a, p. 91). One can make comparisons of structure for example, and endeavour to persuade another that a particular structure is more appealing (1970b, p. 228). One has to admit however, that "the existing separation between the sciences and the arts is artificial" (1970a, p. 92). "Neither Lakatos nor anybody else has shown that science is better than witchcraft and that science proceeds in a rational way. Taste not argument guides our choice of science" (1973). Essentially, whether or not one tries to persuade others of a particular point of view, in the sciences, as in the arts, matters of taste are the final determinants and not matters of rationality.

Objections to Feyerabend's amethodology

Any attack upon Feyerabend could proceed along a number of lines: historically it is ill-supported; logically his methodology gives no grounds for believing that its use guarantees the improvement of knowledge; people do not need to be free in the sense that Feyerabend suggests; a methodology proposing proliferation of theories approves of methodologies other than its own; the theory is inconsistent.

(i) *Ill-supported by history?* The complexities of the scientific enterprise surveyed at any one time or over a period of time, are of such a magnitude to make an objection to his position, on historical grounds, somewhat difficult. There would seem to be such a great wealth of material that one could draw upon to lend support to a score of positions. Lakatos can see Newton working along the lines of sophisticated methodological falsificationism (Lakatos, 1970, pp. 133–138); Kuhn can see paradigm shifts in the 'discovery' of oxygen (Kuhn, 1970a, pp. 52–56, 62–65); while Feyerabend can see Galileo being *ad hoc* and using propaganda in the interests of a conceptual change. It is not difficult to see how Feyerabend could gain support for his belief in the existence of incommensurable theories, and in the importance of principles of proliferation and tenacity at various times in the history of scientific development. When, however, Feyerabend commits himself to an explicit all-inclusive view on the nature of normal science, revolutions, and the proliferation of theories, he makes himself more vulnerable to attack. Kuhn believes that a period of single-paradigm activity is followed by a period of multi-paradigm activity and revolution. Feyerabend believes that normal science and the proliferation of theories co-exist with normal science changing from time to time, for any number of reasons. Who is right? The conclusion one might come to is that the nature of the scientific enterprise is so complex, that neither is right, only a synthesis of a variety of views does justice to the 'reality'.

Whether history supports his claim that an anarchistic epistemology such as his is a better means for improving knowledge, or not, is difficult to tell. It partly depends on what is meant by 'improving knowledge'. It might be that he would argue that the statement is in some senese analytically true: the change in knowledge which follows from an anarchistic epistemology has to be better than that which follows from any other epistemology. One could try to argue that the incommensurable theories which Feyerabend places such weight upon are relatively rare, and consequently should not feature significantly in a description of scientific methodology as it is actually practised. One could also challenge such universal statements as: 'no theory ever agrees with all the facts', or 'all rules are violated at some time or another, and such violations are necessary for progress'. Of course one of the fundamental problems in opposing a view like Feyerabend's, on historical grounds, is that Feyerabend would see the interplay and clash of various epistemological theories as an inevitable result.

Ultimately, however, his principle 'anything goes' is not open to attack on historical grounds. It is fundamentally a prescriptive methodology and not descriptive, although Feyerabend does claim that the justification of the principle resides in 'the advance of science' (1973).

(ii) *Improvement of knowledge is not guaranteed by his methodology?* Little attention is given by Feyerabend to the notion of scientific progress. As mentioned above, he does claim that his anarchistic epistemology as opposed to other epistemologies is a better means for improving knowledge (1970a, p. 21). He also asserts,that violations of methodological rules is necessary for progress (1970a, p. 22). Again, he proposes that knowledge is best improved if one aims at maximum testability of knowledge (1965, pp. 223–225). Nowhere, however, does Feyerabend tell us what

he means by 'progress' or 'improvement' of knowledge'. It may be that this notion of maximum testability accounts for the progress that Feyerabend still thinks that science makes – 'the advance of science' is the touchstone by which to judge the failure of all methodologies and to justify 'anything goes' (1973) – and perhaps we are to assume that advancement in science is equivalent to improvement in knowledge. It is difficult to say. But granted that this advancement and/or 'improvement' has anything to do with anybody's decision that progress has been made, regardless of what basis they use for this decision, it is difficult to see how the principle 'anything goes' guarantees that such progress is made. It cannot simply be that progress is made as theories are proliferated, for how then can the progress made justify the principle? Progress is in that case identified with the out-working of the principle. Of course, it is difficult to see how any principle, without progress written into it, could guarantee progress.

Perhaps progress can be guaranteed only in the sense that as the scientific communities make decisions, they will not be prepared to say that anything other than progress has been made (Kuhn 1970a, p. 166) – a sociological and a psychological phenomenon. It is another matter to argue that scientific communities tend only to make rational decisions. It now becomes an historical problem, and Feyerabend, as has been argued above, is relatively immune to many historical attacks.

(iii) *People do not need to be free?* Feyerabend has made it very clear that one of the main reasons for propagating his principle 'anything goes' is that it is only such a principle which is consistent with a humanitarian outlook and which is appropriate for free men. Subscription to methodological rules enslaves people to rigid procedures. Now obviously, Feyerabend is making an appeal, using, for many, the attractive notions of 'freedom' and 'humanitarianism'. It is questionable whether people need to be 'free' or even want to be free in the sense of not being bound by methodological rules. Feyerabend would undoubtedly say that for their own sakes they do need to be free in that sense. There seems to be no *easy* way of counteracting the claim or of giving rational support to the claim. In a debate of this kind, about the 'proper' role for man in his world, disputants on both sides of the argument are left holding their various metaphysical beliefs. Of course one could try to argue at a psychological or sociological level, but metaphysical beliefs would still not be absent in such a debate.

(iv) *Other methodologies are approved by Feyerabend's methodology?* If one takes Feyerabend's notions of 'proliferation of theories' and 'maximum testability' seriously, then Feyerabend could be expected to suggest that in order to best test his own epistemological theory, many other epistemological theories [presumably of various grades between methodological and amethodological and many types of methodological theories (and amethodological?)] should be invented. He does not suggest this. Perhaps he regards his theory as simply adding to the many theories already in existence.

Alternatively, Feyerabend may simply deny the applicability of his principle to theories about theories. 'Anything goes' may be a principle in a theory about knowledge, but he may hold as another metaphysical principle that it is not applicable in a theory about theories of knowledge. However, so insistent is he upon man being free of methodological rules, that it would be understandable if he argues that theories of knowledge should abound except for those that entailed methodological oppression.

(v) *The theory is inconsistent?* There does appear to be some inconsistency in his 'methodology'. While the salient feature of it is the principle 'anything goes', and expositions of that principle

suggest that that means *anything* goes, it seems that none the less Feyerabend assumes that by the use of the principle progress is made. Some theories are in the end eliminated (though not necessarily permanently). Whatever might be meant by progress, progress is presumably made because some decisions are taken as to the abandonment of theories. Surely however, the principle 'anything goes' means not only that theories can be resurrected but also that no theories need ever be abandoned – let the principle of proliferation *and* the principle of tenacity have their full force. Because of this inconsistency, and because there is no prescription for the abandonment of theories (there cannot be for the methodological-free man), there exist the nagging questions: what is meant by progress? how could we recognise it? how could we contribute to it? is it really the case that Feyerabend doesn't want to prescribe for us some methodological rule for the abandonment of theories, no matter how vague?

Feyerabend and science education

To the extent to which teaching about science implies teaching about actual scientific methodology, in the interest of validity, a realistic presentation of scientific methodology is required.

If Feyerabend is 'right' then what most science teachers teach as methodology in science is almost certainly in many respects false. Their views and what they pass on to others are too simplistic and too much characterised by methodological rules. One wonders what place is given in the teaching of science to the notions of 'normal science', 'proliferation of theories', their interplay, and 'scientific revolutions'. Where would students be taught that the methodological rule 'anything goes', with its consequent principles of proliferation and tenacity, is the only appropriate rule for the advancement of science? How much attention is given by teachers of science to the theory laden character of so-called facts, concept: fact mobility, and the use of propaganda in persuading people to change from one theory to another? To what extent are students made familiar with the existence of incommensurable theories and the limitations of logic and rationality in decision making in science?

Of course, it must not be thought that only Feyerabend has argued for all these points of view. Popper and Lakatos are in agreement with him on some matters, and Kuhn agrees with him on a number of issues. But the total view alluded to above is distinctly Feyerabend's. One cannot be Popperian, Kuhnian or a disciple of Lakatos and adopt this view of scientific methodology, only certain aspects of it. If one is committed to Feyerabend and the teaching of scientific methodology, then one is committed to the teaching of the views outlined above.

It may be however, that only certain of these notions appeal. Perhaps the idea – the only principle that is proper for free persons and that is appropriate for the advancement of science is the principle 'anything goes' – seems too outrageous. Perhaps some would settle for something less, such as, 'anything could go' and 'sometimes it does'. Perhaps some would opt for a synthesis of a number of views for a better 'match' for reality, and for a better chance for progress. Whatever the case, our ideas on scientific methodology and consequently our ideas on teaching scientific methodology may well change after a consideration of the Feyerabendian position.

Feyerabend and science education research

If science teachers portray an improper picture of scientific methodology, according to Feyerabend, science teacher educationists operate by one. Feyerabend would suggest that we recognise and question the theories behind our 'facts', that we consider the possibility of incom-

mensurable theories as we argue for theory preference, that we abandon theories only with reluctance after we have generated other theories, all sorts of theories, and that we appreciate the role of propaganda and the limitations of logic and rationality in our efforts to advance.

The opening address of the Fifth Annual Conference of the Australian Science Education Research Association ended with a quotation from Feyerabend (1968 in Phillips, 1974) – a statement pointing out the desirability of speculation and the invention of alternatives. Phillips suggested that as a first step towards liberating educational research from the conservative influences that restrict it, researchers should try and recognise the value of such speculation and novel approaches. It is, simply, a first step. To follow Feyerabend further one needs not only to see the value of novelty, nor even to try one's hand at new approaches, but to consider the desirability of at least sometimes working on the principle 'anything goes'. The last step, doubtless not to be embraced by all, is to be committed *always* to such a principle.

Note

1. Much of what follows is taken from the author's unpublished B.A. thesis (Sydney University, 1973).

References

BERKSON, W. "Some practical issues in the recent controversy on the nature of scientific revolutions." *Boston Studies in the Philosophy of Science*, **14**: 197-210, 1974.

CHRISTIAN, D. "Verification and falsification in science." *The Australian Science Teachers Journal*, **19**(3): 33-40, 1973.

COLLETTE, A.T. *Science Teaching in the Secondary School*, Allyn and Bacon, Boston, 1973.

FEYERABEND, P.K. "Explanation, reduction and empiricism." *Minnesota Studies in the Philosophy of Science*, **3**: 28-97, 1962.

FEYERABEND, P.K. "Reply to criticism." *Boston Studies in the Philosophy of Science*, **2**: 223-261, 1965.

FEYERABEND, P.K. "How to be a good empiricist", in P.H. Nidditch (ed.), *The Philosophy of Science*. University Press, Oxford, 1968.

FEYERABEND, P.K. "Against method." *Minnesota Studies in the Philosophy of Science*, **4**: 17-130, 1970(a).

FEYERABEND, P.K. "Consolations for the specialist", in I. Lakatos & A. Musgrave (eds.), *Criticism and the Growth of Knowledge*, University Press, Cambridge, 1970(b).

FEYERABEND, P.K. "Problems of empiricism, Part II", in R.G. Colodny (ed.), *The Nature and Function of Scientific Theories*, University of Pittsburg Press, 1970(c).

FEYERABEND, P.K. "Theses on anarchism", Mimeo by Paul Feyerabend, for a debate with Imre Lakatos at Boston Colloquim, 20th March, 1973.

HEMPEL, C.G. & OPPENHEIM, P. "Studies in the logic of explanation." *Philosophy of Science*, **15**: 135-175, 1948.

HURD, P. *New Directions in Teaching Secondary School Science*. Rand McNally, Chicago, 1969.

JEVONS, F.R. *The Teaching of Science*. Allen and Unwin, London, 1969.

KUHN, T.S. *The Structure of Scientific Revolutions*. University of Chicago Press, 2nd. ed., 1970(a).

KUHN, T.S. "Reflections on my critics", in I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge*. University Press, Cambridge, 1970(b).

LAKATOS, I. "Falsification and the methodology of scientific research programmes", in I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge*. University Press, Cambridge, 1970.

LAKATOS, I. "The role of crucial experiments in science." *Studies in the History and Philosophy of Science*, **4**(*4*): 309-325, 1974.

LAKATOS, I. & MUSGRAVE, A. *Criticism and the Growth of Knowledge*. University Press, Cambridge, 1970.

MACKLIN, M.J. "On hiding ones curriculum under an ideological bushel." *The Australian Science Teachers Journal*, **19**(*3*): 53-65, 1973.

MARTIN, M. *Concepts of Science Education*. Scott, Foresman, Glenview, 1972.

MAXWELL, N. "A critique of Popper's views on scientific method." *Philosophy of Science*, **39**(*2*): 131-152, 1972.

NAGEL, E. "The meaning of reduction in the natural sciences", in R.C. Stauffer (ed.), *Science and Civilisation*. University of Wisconsin Press, Madison, 1949.

NOVAKOVIC, S. "Is the transition from an old theory to a new one of a sudden and unexpected character?" *Boston Studies in the Philosophy of Science*, **14**: 173-196, 1974.

PHILLIPS, D.C. "Can scientific method be taught?" *The Australian Science Teachers Journal*, **14**(*2*): 32-41, 1968.

PHILLIPS' D.C. "Conservative influences on experimental research." *Research in Science Education*, **4**: 3-10, 1974.

POPPER, K.R. *The Logic of Scientific Discovery*. Hutchinson, London, 6th impression (revised), 1972(a).

POPPER, K.R. *Objective Knowledge*. Oxford University Press, London, 1972(b).

QUINN, P. "Methodological appraisal and heuristic advice." *Studies in the History and Philosophy of Science*, **3**(*2*): 135-149, 1972.

RENNER, J.W. & STAFFORD, D.G. *Teaching Science in the Secondary School*. Harper and Row, New York, 1972.

SUND, R.B. & TROWBRIDGE, L.W. *Teaching Science by Inquiry in the Secondary School*. Merrill, Columbus, Ohio, 2nd ed., 1973.

TAWNEY, D.A. "The nature of science and scientific enquiry", in C.R. Sutton & J.T. Haysom (eds.), *The Art of the Science Teacher*. McGraw-Hill, London, 1974.

TISHER, R.P., POWER, C.N. & ENDEAN, L. *Fundamental Issues in Science Education*. Wiley, Sydney, 1972.

TRICKER, R.A.R. *The Contribution of Science to Education*. Mills and Boon, London, 1967.

YOUNG, R.E. "Research as a knowledge process: A critical view." *Research in Science Education*, **4**: 199-203, 1974.

VALIDITY OF PERCEIVED NEED SATISFACTION AS A CURRICULUM EVALUATION CRITERION

Lindsay D. Mackay & Jeff R. Northfield

Purpose

The purpose of this paper is to report on investigations of the validity of one of the criteria being used in the formative evaluation of ASTEP materials: the extent to which the materials are perceived to satisfy certain student-teacher 'needs'. Before embarking on a description of the results, some of the background to the use of this criterion will be outlined.

Background: Satisfaction of perceived student-teacher needs as a criterion for STEP materials evaluation

Student teacher motivation has been an important consideration in the development of materials for the Australian Science Teacher Education Project (ASTEP) and for its United Kingdom predecessor, the Science Teacher Education Project (STEP). The case for student motivation as a component in development of curriculum innovations and as a criterion for curriculum evaluation has been argued elsewhere (Haysom and Sutton, 1973/74, 1974), and will not be repeated in detail here. In an attempt to establish motivational criteria for the evaluation of STEP materials, Haysom and Sutton derived a list of statements of student needs which they described as:

> characteristics of the learner which can be harnessed to encourage a search for goals and objectives which might not have been immediately obvious to him.
>
> (Haysom and Sutton 1973/74, p. 28)

This list was derived from tutor's reports on the reasons that student-teachers responded to, and became involved in, STEP activities, and from tape recordings of interviews with student-teachers on their reactions to STEP activities they had experienced. The way in which the list of statements was derived is described in detail elsewhere (Science Teacher Education Project, 1971; Haysom and Sutton, 1974).

The uses to which these statements of perceived student-teacher needs were put in the evaluation of STEP materials is described in detail by Holford (1974). Basically, student-teachers who had experienced a STEP activity were asked to rate their reaction to an activity and indicate whether or not each of the eleven needs statements were reasons that the activity appealed to them. Tutors were also asked to rate the extent to which the unit motivated student-teachers, and to indicate which of the list of eleven needs statements were met by the materials. This information from students and tutors was used in assessing the effectiveness of units.

The use of self perceived needs of student-teachers in the evaluation of ASTEP materials

Student needs have also occupied a central role in the development and evaluation of the ASTEP materials (Fensham, Northfield & Driscoll, 1974; Mackay, 1973a; Mackay & Northfield, 1974). Initial research conducted with the needs statements in Australia (Mackay, 1973b; Gunstone & Mackay, 1975) has indicated:

(a) that all but one of the eleven initial statements of needs are consistently

interpreted, as far as meaning is concerned, by Australian student-teachers;

(b) that the needs statements are ranked in a similar order of importance by Australian and British student-teachers despite the diversity in attributes of the two groups (correlation between median ranking of statements by the two groups = 0.6); and

(c) that the perceived importance of the need statements changes during a course of teacher-training.

On the basis of the experience of STEP and the acceptance of the STEP philosophy that student teacher motivation and perceived satisfaction of the felt needs of student-teachers were important objectives for materials to be used in teacher education courses, perceived needs satisfaction was included as one of the criteria to be employed in the formative evaluation of ASTEP materials.

Data was collected from four sources:

(a) the author(s) of the unit were asked to specify the *student needs which are met* by that unit;

(b) student-teachers who had experienced an ASTEP unit were asked to indicate the extent to which they consider the unit contributed to the satisfaction of needs expressed in a number of statements on the three-point scale:

not at all or to a negligible extent,
to some extent,
to a substantial extent;

(c) tutors who had used the unit in their course were asked to indicate the extent to which they considered the unit *contributed to the satisfaction of student needs* expressed in a number of statements on the same three-point scale;

(d) armchair critics who conducted an armchair evaluation of a unit were asked to indicate the extent to which they consider the unit *is likely to contribute to the satisfaction of student needs* expressed in a number of statements.

Of the lists of statements used to gather data from these four sources, ten statements were common to all four lists of need statements. These were:

1. To feel that you will cope more effectively in the classroom.
2. To feel you have achieved success (by your own standards).
3. To have new or fresh experience.
4. To sort out or express your own ideas or opinions.
5. To work with others.
6. To feel that your contribution is appreciated by others.
7. To see yourself more clearly as a teacher.
8. To feel you have improved your practical teaching.
9. To feel you have developed a better understanding of children.
10. To feel you have developed a better understanding of science.

Information on the extent to which the four groups of observers (authors, student-teachers, armchair evaluators and tutors) perceived that these common statements of student-teacher needs were being satisfied, or were likely to be satisfied, by ASTEP units form the basis of this investigation.

Validity of perceived needs satisfaction as a criterion for materials evaluation : the questions

The validity of a criterion such as perceived needs satisfaction cannot be directly measured and expressed as a single index in the same way as the predictive validity of an HSC examination can be expressed in terms of its correlation with tertiary performance. Rather, the validity must be established in terms of (i) the extent to which the information on perceived needs satisfaction is internally consistent and (ii) the extent to which the information is consistent with the initial construct of needs-satisfaction within the student motivation model from which its use is derived.

Four questions were posed in an attempt to investigate the validity of perceived needs satisfaction as a criterion.

Question 1. Are various ASTEP units perceived to satisfy different needs? If the criterion does not distinguish between units, then its usefulness as a criterion is limited. Figure 1 illustrates the needs perceived to be satisfied by student-teachers for four units with at least forty student-teachers returns from at least two institutions. In this figure, the eleven needs statements are grouped into four groups of needs based on previous analyses.

Personal Needs:

4. To sort out or express your own ideas or opinions.
3. To have new or fresh experience.
2. To feel you have achieved success (by your own standards).

Social Needs:

5. To work with others.
6. To feel that your contribution is appreciated by others.

Coping as a Teacher Needs:

1. To feel that you will cope more effectively in the classroom.
8. To feel you have improved your practical teaching.
7. To see yourself more clearly as a teacher.

Understanding Needs:

9. To feel you have developed a better understanding of children.
10. To feel you have developed a better understanding of science.
11. To feel you have developed a better understanding of science education.

It is apparent from Figure 1 that different patterns of needs are perceived to be satisfied by various units. Thus the criterion of perceived needs satisfaction does discriminate between ASTEP units and can be used to distinguish between units.

Question 2. Are the needs perceived by student-teachers to be satisfied by ASTEP units those which the authors specify as met by the units? An answer to this question can be obtained by comparing the mean percentage of student-teachers who perceive that a need is satisfied by a unit for two sets of needs –

- those nominated by the authors as met by the unit; and
- the other needs.

One would anticipate that a higher percentage of student-teachers would indicate that the needs

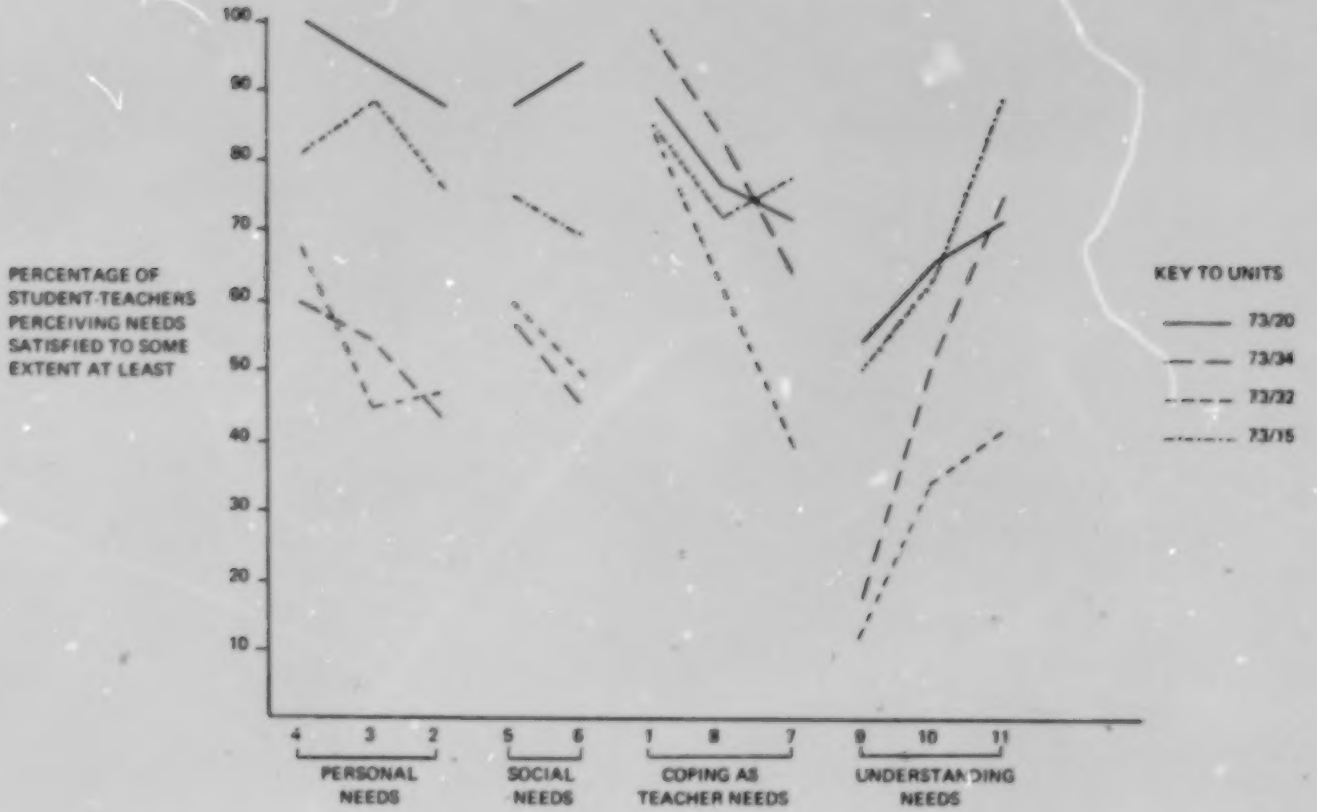


FIGURE 1: Needs perceived to be satisfied by four ASTEP units

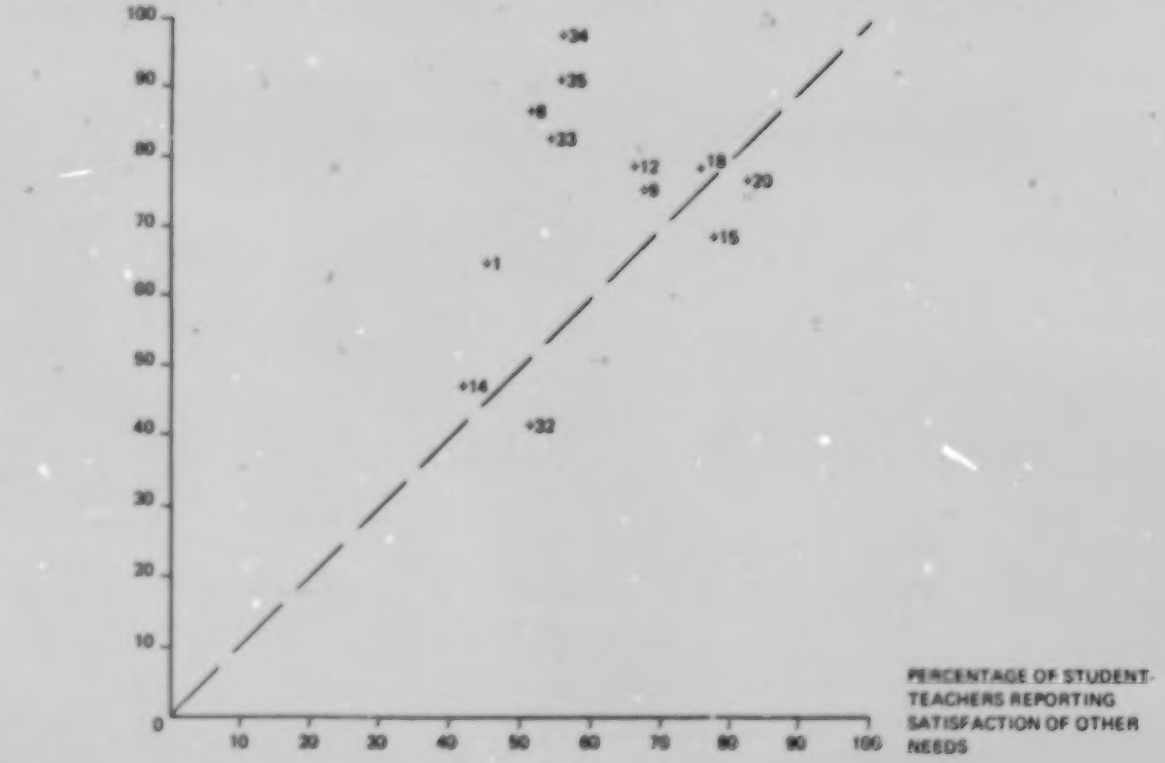


FIGURE 2: Consistency in needs perceived as satisfied by student teachers and those nominated as met by authors.

nominated by the authors were satisfied by the unit than would be the case for other needs. This is so for nine of the twelve units with more than ten student-teacher evaluation sheets available as indicated in Figure 2. For five of the units a Mann-Whitney U test indicates that the difference in the anticipated direction is significant beyond the $\alpha = .05$ level. A Kolmogorov-Smirnov test applied to the values of U obtained for the whole set of twelve units indicates that there is a tendency for needs nominated by authors as met to be more frequently reported by student teachers to be satisfied by the unit than for other needs.

Question 3. Do student-teachers, armchair evaluators and tutors consistently report a unit as satisfying the same needs? This analysis was restricted to analysis of the data on six units which had a satisfactory amount of data for two or three groups of reporters on the unit, and for which the data happened to be readily available at Monash.

In order to answer this question, Spearman Rank-order correlation coefficients were computed for each unit between the percentage of respondents reporting a need as satisfied for various groups of reporters. The results are summarised in Table 1.

TABLE 1

Spearman Rank-order correlations between percentage of respondents reporting various needs as satisfied by an ASTEP unit for three groups of respondents

Unit	Correlation between student-teachers and armchair critics	Correlation between student-teachers and tutors	Correlation between armchair critics and tutors
73/05	.50	.33	.68
73/08	.41	.14	.61
73/34	.06	.83	.37
73/35	.39	.44	.86
73/01	.74	-	-
73/07	-	.76	-
Median value	.41	.44	.64

For all three comparisons of pairs of respondents in Table 1, the set of correlations is such that a Kolmogorov-Smirnov test indicates that there is a significant correlation (beyond $\alpha = .05$) between the percentages of respondents in the three groups reporting needs as satisfied by ASTEP units.

Question 4. Is a student's overall rating of a unit related to his/her perceptions of the extent to which the unit contributes to satisfaction of specified needs? At the simplest level, one can test the proposition that the more needs a student-teacher perceives as satisfied, the more highly he or she will rate the unit. This corresponds to the "needs-richness" notion proposed by Haysom & Sutton (1974).

The data for 158 student-teachers who reported on 10 of the ASTEP units is summarised in Figure 3. There is a positive correlation ($r = .12$, $p < .07$) between number of needs reported as

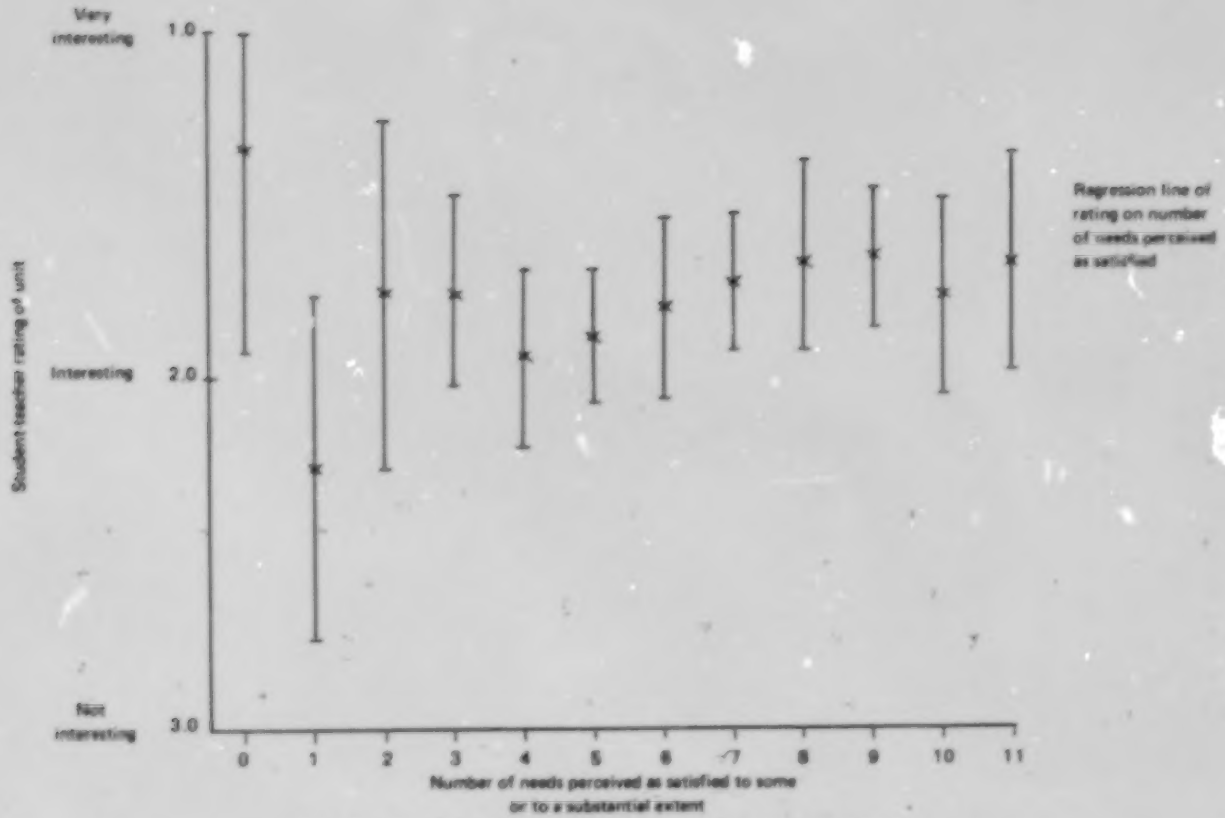


FIGURE 3: Relationship between student-teacher rating of a unit and number of needs perceived to be satisfied

satisfied and how interesting the student-teacher found the unit. This indicates that the number of needs a student-teacher reported as satisfied by a unit could account for about 1.4% of variance in his or her rating of the unit.

At a more sophisticated level, one can establish a model which allows for different contributions to the rating of a unit for satisfaction of different needs and for different levels of perceived satisfaction (to some extent and to a substantial extent). If one carries out an analysis based on such a model, 19% of the variance in student-teacher ratings of a unit can be accounted for.

As anticipated, when a need was perceived as satisfied to a substantial extent it made a greater mean contribution to rating of the unit. In addition, if one takes the mean contribution that satisfaction of a need makes to the rating of a unit as an indication of the perceived importance of that need, this can be compared to the data on the reported importance of these needs for a large sample of student-teachers obtained by Gunstone and Mackay (1975). There was a Spearman rank correlation of .738 between the reported importance of these needs statements and the mean contribution that perceived satisfaction to a substantial extent of this need made to the rating of that unit.

Summary and conclusion

In summary, then, the results reported in this paper provide clear support for the validity of perceived need satisfaction as a criterion for evaluation of ASTEP materials.

Specifically it has been demonstrated that:

- different ASTEP units are perceived to satisfy different student-teacher needs;
- authors of ASTEP units, student-teachers, tutors and armchair critics exhibit consistent perceptions of the needs satisfied by particular ASTEP units; and
- perceived satisfaction of needs is related to the student-teacher's rating of that unit in terms of its interest, and the more important the need is perceived to be, the greater the contribution that its satisfaction appears to make to the rating of the unit.

Having demonstrated that perceived needs satisfaction is a valid criterion for the evaluation of teacher education materials, this raises the question of whether perceived needs satisfaction could or should be a criterion for evaluation of other curriculum materials.

References

FENSHAM, P., NORTHFIELD, J. & DRISCOLL, D. "ASTEP – A model for developing curriculum materials in teacher education." *South Pacific Journal of Teacher Education*, **2**(*1*): 5-11, 1974.

GUNSTONE, R.F. & MACKAY, L.D. "The self perceived needs of student teachers." *South Pacific Journal of Teacher Education*, **3**(*1*): 44-51, 1975.

HAYSOM, J.T. & SUTTON, C.R. "Motivation: a neglected component in models for curriculum improvement." *Curriculum Theory Network*, **4**(*1*): 23-35, 1973/4.

HAYSOM, J.T. & SUTTON, C.R. *Innovation in Teacher Education*. McGraw Hill, London, 1974.

HOLFORD, D.G. "Evaluation: A process of distilling experience" and "Case studies in the evaluation of particular curriculum units", in Haysom, J.T. and Sutton, C.R. *Innovation in Teacher Education*. McGraw Hill, London, 1974, 77-90 and 94-109.

MACKAY, L.D. "A proposal for the formative evaluation of ASTEP materials." Mimeographed Report, ASTEP, Canberra, 1973(a).

MACKAY, L.D. "ASTEP interim research report no. 1/73." Mimeographed Report, ASTEP, Canberra, 1973(b).

MACKAY, L.D. & NORTHFIELD, J.R. "The use of a formative evaluation model for ASTEP." *Research in Science Education*, **4**: 65-75, 1974.

SCIENCE TEACHER EDUCATION PROJECT. "With student teachers in mind." *Information Bulletin No. 3*, 1971 (mimeo).

SCIENCE EDUCATION & THE SUPERNATURAL IN PAPUA NEW GUINEA

R.J. Stanton & M.N. Maddock

Background

Science teaching has been an essential part of formal education in high schools in Papua New Guinea since about 1960, and with the implementation of the Three-Phase Primary Science program in primary schools in the early 1970's, now forms part of the total school experience of all pupils passing through the school system, whether their formal education ends at standard 6 primary school, or at post secondary level.

Both the authors have been closely associated with science teaching and science curriculum development in the country since 1963, and both have developed deep interest in and concern for the relationships between the impact of formal schooling, with its science teaching component, and the traditional culture of the Papua New Guinean people.

This study represents a culmination of this interest, and looks at the attitudes of Papua New Guineans in the area of their views of natural phenomena and the realm of the supernatural.

Papua New Guinea is a country of about two million people speaking about 700 different languages, with a wide variation of traditional culture patterns, but with a common dimension in belief in the use of magic and sorcery, and in the existence of spirit beings.

In most Melanesian belief systems, the realm of the non-empirical is always closely associated with, and in most cases part of the ordinary physical world – it is supernatural only in a limited sense (Lawrence and Meggitt, 1965, p. 9). The gods, ghosts, ancestors, demons and totems are generally said to live on earth, or in a world which is a physical replica of the earth. Although there are exceptions, very many of the peoples believe in similar types of spirit beings, some of which are regarded as malevolent. For example, the Mae Enga attribute most injuries, illnesses and deaths to evil spirits (Meggitt, 1965) and the Lalakai of New Britain believe in malevolent bush spirits which go to much greater extremes than human beings in their reactions: some are malevolent without apparent cause and kill off, or drive away, whole villages of men (Valentine, 1965).

Magic and sorcery form complex dimensions within the cultures, complicating attitudes towards natural phenomena, the environment, the ecological cycle, and bodily health and well being. Magic is resorted to when the observational and reasoning skills accomplished by the culture cannot be reliably used to account for natural phenomena; magic is related to the "domain of the unaccountable", (Malinowski, 1948).

A formal education system which ignores such elements in the culture of the people in its curriculum development program is asking for trouble. A science curriculum which presents science in the context of western culture to Papua New Guinean students and does not take any account of local attitudes and beliefs runs the risk of cultural conflicts and the development of tensions.

The stated aims of the Papua New Guinea science syllabus emphasise awareness of, interest in and curiosity about natural phenomena, and a commitment to seek scientific explanation of these phenomena; understanding and appreciation of a person's relationship with, and confidence in ability to affect changes and improvements in, his environment; and critical thinking ability and reduction in tendency to adopt opinions based on unsupported or unreliable evidence, (Department of Education Papua New Guinea 1971). To date, Papua New Guinea's science syllabuses

have presented science to students in a western cultural context, with little or no account of the fact that a Papua New Guinean's view of his relationship with natural phenomena and the environment may adopt quite a different stance.

The importance of science education taking account of traditional attitudes has been strongly advocated by UNESCO in a report by the Advisory Committee on Application of Science and Technology to development. Science education must include an examination of the extent to which the traditional attitudes, basic assumptions and concepts are accepted and used by the students in every-day life; the traditional world response may be the source of particular difficulties for school children in their learning of science (UNESCO, 1970, para. 17).

Stone (1967) expressed concern about the introduction to Africa, which has many similarities to the Papua New Guinea situation, of curriculum projects from the United States of America and the United Kingdom. These projects, developed for and within the most western of societies, assume rational and mechanistic terms of reference. Stone emphasised the need for research into African intellectual backgrounds.

Stone's research with the Yoruba people of Nigeria showed that there was a general tendency towards a western secular approach between the beginning and end of secondary schooling, and that non-secondary school teenagers remained markedly traditional in their responses in terms of the assumptions and concepts they accept and use in life.

Science has largely ignored the traditional world of Papua New Guineans, putting forward its own view regardless, but many religious bodies have placed themselves in direct conflict with it, particularly in regard to belief in spirits. Both authors have heard teachers provided by various religious denominations impressing upon their pupils in the weekly period of Religious Instruction that spirits do not exist, and in response to the question "do spirits exist?", having the class chorus the response "no, spirits do not exist". With pressures on traditional views from sources such as science and religion, it would be reasonable to assume that some form of dualism would become evident.

Dart and Pradhan (1967) made a study of attitudes towards familiar phenomena of nature and sources of knowledge about nature in 9-14 year old children in Nepal and found evidence of dualism, with individual children and groups of children giving both folk-oriented and school oriented explanations. Amongst outsiders working in Papua New Guinea, it has often been said that the local people exhibit a similar attitudinal dichotomy, but recent research efforts have not produced any concrete evidence of such dualism.

Prince (1969) sought the views of Papua New Guinean students on physical causality, investigating explanations about rain, thunder, lightning, eclipses, earthquakes and sickness. Ninety eight percent of responses represented at least an attempt at Western causal explanation, with no evidence of dualistic attitudes showing through.

However there has been very little specific research done on the effect of secondary education on the cultural development of students in Papua New Guinea.

Previous research as the foundation for this study

Evil spirit scale. In an attempt to gain information on the belief held by high school students in the physical powers of "evil spirits" Stanton used the paired comparison technique (Edwards, 1957) in 1969 to develop a preliminary form of an instrument aimed at providing information on the belief held by Papua New Guinean students in the physical powers of "evil spirits" which could cause death. The study has not been published.

A folk-type story with strong parallels with Papua New Guinean folk ideas was set in South America. It was about an Indian man who went into a cave, which the local people believed was inhabited by evil spirits, to shoot some flying foxes to eat. The gun fired, the ground shook, the man screamed and was never seen again. The South American setting was chosen to give a degree of detatchment to the situation, as in Stanton's experience direct questioning concerning spirits results in direct denial of any commitment to such a belief.

Eleven items stating possible causes for the disappearance of the man were used from an original pool of 19 items; only one of those chosen had definite supernatural implications. These were paired to produce 55 paired comparisons, 10 of which matched "evil spirit" with other explanations. The possibilities ranged from very likely natural ones through to almost impossible.

The instrument was constructed in the same format as a pencil and paper English comprehension test with which the students were familiar, and the first pair involving the "evil spirit" was item 10, allowing the subject to move comfortably into the exercise before confrontation with a "supernatural" possibility. In this way it was hoped that the subjects would react as if in commentary on another culture, rather than being confronted with their own beliefs. The language was pitched at a simple level, comprehensible by all levels of high school students in Papua New Guinea.

The scale was used on a sample of 80 students at a highland high school in Papua New Guinea in 1969, 20 from each of Forms 1, 2, 3, and 4. As the study was a simple, preliminary study, no attempt was made to use a more comprehensive sample, and thus there are limitations imposed on the interpretation of the results.

Using the method as presented by Edwards (1957), raw scores were converted to decimal proportion scores, normal deviates obtained and averaged and an arbitrary zero item chosen. The test for the case V model was applied to each of the school forms and supported the underlying assumptions, although similar support was not obtained for the scale across the four forms. To obtain scores for individual students, each time the "evil spirit" alternative was chosen, the scale value of the item over which it was selected was allocated as the score. Those items not involving evil spirits were scored zero. The procedure was adopted because it was considered that selection of the "evil spirit" cause over a strongly logical natural explanation was stronger indication of strength of attitude than selection over a relatively impossible alternative.

Although this preliminary study indicated problems with some of the items, the scale discriminated between individuals on the basis of scores, and examination of the evil-spirit items showed that for the sample, great credence was placed on the possibility of an "evil spirit" causing the death or disappearance of a man by students in Forms 1 to 3, with a considerable change of pattern in Form 4.

Environmental phenomena attitude scale. During 1972 Maddock conducted the field work for a study (Maddock, 1973) in which an instrument called the Environmental Phenomena Attitude Scale was developed to measure the attitude of Papua New Guineans to investigation, control and manipulation of natural phenomena. Various aspects of this study are reported in Maddock (1973), Maddock (1974a, 1974b) and Maddock (1975a, 1975b).

A reliable scale of 26 items was produced in each of which a concrete situation related to Papua New Guinean culture was described and three choices of though or action given. The subject was asked to select the alternative most closely aligned with his own way of thinking or acting. These were scored on a 3 point scale, with the 3 points allocated to the alternative reflecting the positive end of the continuum. The instrument was administered to a broad-spectrum sample of

Papua New Guineans covering regional and ethnic dimensions, urban and rural dimensions, and matching third and fourth form high school students with people of about the same age, speaking the same language, but with minimal formal schooling.

Face to face interviews were conducted in the English, Pidgin, Hiri Motu and Enga languages. The instrument was also administered to a larger sample of students using a group-interview technique in which the questions were read out aloud to a group, but the answers marked on an answer sheet by the individuals. In the face to face interviews the students were asked to predict answers they expected uneducated villagers to give. In summary, a marked significant difference was found in attitude, as measured by the scale, between the students and minimally educated villagers, but no regional, life-style, sex or ethnic differences were detected. Students predicted a wide gap between their own attitude and the attitudes held by uneducated villagers, but the actual gap determined was much smaller than the students predicted. In open-ended interview following the administration of the instrument, the students forcibly expressed a patronizing view of village culture and attributed their own change of attitude from that held by villagers to the influence of education, with frequent reference to science lessons and experiments.

Although no specific reference was made to "spirits" in the concrete situations described in the items, certain items could be interpreted as having supernatural implications by subjects with strong beliefs in the existence of spirits. Two items in particular, one concerning the influence of a cave, and another concerning a forbidden mountain, could readily be interpreted in this fashion.

Rationale, sample and design

During 1974, Maddock carried out a follow-up project using the Environmental Phenomena Attitude scale on a research grant from the Education Department of Papua New Guinea. The larger sample covered the New Guinea Islands as well as the mainland; contained a range of educational levels from no formal schooling through to tertiary level; had an age range from about twelve years of age to elderly, and used three additional languages. As an adjunct to it it was decided to conduct a preliminary small scale study using Stanton's Evil Spirit Scale and Maddock's Environmental Phenomena Attitude Scale in conjunction to investigate relationships between the two attitudes. The study provided an opportunity to look at the possibility of dualistic attitudes existing. Neither Prince (1969) in his investigation of the views of Papua New Guinean students on physical causality nor Maddock (1973) were able to detect dualistic attitudes.

To internalize acceptance of natural "cause and effect", as understood in the context of formalized science, one might assume that there must be a certain degree of attitude change in which a person might have to reject inadequate concepts and related attitudes generated within his culture. Such a change could create traumatic problems which might in some way be alleviated by some form of dualism by setting up an additudinal dichotomy when faced with the dilemma of conciliating conflicting cultural demands. It would thus be possible to survive well in a superimposed cultural milieu, while still referencing behaviour to the native custom. The results of the Maddock (1973) study showed that a strong difference of attitude exists between the high school students and their village counterparts, and that the students are acutely aware of this difference.

The Evil Spirit Scale and the Environmental Phenomena Scale were administered to a sample of 103 Papua New Guinean High School students, drawn from Forms 1 to 4 at three high schools – two in the Southern Highlands and one in East New Britain. Selection of members of the sample was done in as random a fashion as possible and, based on Maddock's (1973) findings

of a lack of regional or ethnic effects on scores on the Environmental Phenomena Attitudes scale, can be reasonably considered as representative of Papua New Guinean high school students.

The Evil Spirit Scale was administered in the same manner as in the original study, as a paper and pencil exercise in the fashion of a comprehension test in English. The Environmental Phenomena Attitude Scale was administered using the group-interview technique used in the original study and the subjects were also asked to predict the answers their uneducated village counterparts would give.

Three scores were thus obtained for each subject. A fourth score was obtained by subtracting the prediction score from the student's own score in the case of the Environmental Phenomena Attitude Scale. This score has been called the *Alienation Score* as it is considered to indicate the degree of alienation from the traditional culture in respect to the attitude measured by the scale.

For each of the sets of scores, the sample was investigated for differences on regional, educational level and sex dimensions.

Results

Score reliability. The items on the Evil Spirit Scale were divided into two sets carrying equal score weightings by randomly allocating each pair of scores with the same or very close weightings. The split half correlation was 0.70, significant at the 0.01 level, and yielded a reliability coefficient of 0.82 (Nunnally, 1967, p. 194).

Coefficient α for the student's own scores on the Environmental Pheonomena Attitude Scale was 0.56, which compares with 0.53 obtained for the group interviews on students in the original study (Maddock, 1973), and for the predictions was 0.75, compared with 0.78 obtained from similar scores in face to face interviews in the 1973 project. Details of ranges, means, standard deviations and reliability are given in Table 1.

TABLE 1

Descriptive Statistics for the four scales used

	Maximum Score	Minimum Score	Range	Mean	Standard Deviation	Reliability Coefficient
Evil Spirits	55	0	55	27.9	16.2	0.82
EPAS (Student Scores)	76	53	23	65.1	5.1	0.56
EPAS (Predictions)	62	29	33	45.1	7.3	0.75
Alienation	49	0	49	20.3	11.50	-

Intercorrelations. Intercorrelations obtained between the Evil Spirit Scale, the student's own scores on the Environmental Phenomena Attitude Scale, the Prediction Scores and the Alienation Scores are given in table 2.

Correlation between the Evil Spirit Scale scores and the Environmental Phenomena Attitude Scale scores was small, negative, and not significant at the 0.05 level. This shows there is virtually no relationship between attitude to evil spirits and the attitude measured by the other scale.

TABLE 2

Intercorrelation between scales

	EPAS Predictions for Villagers	Alienation Scores	Evil Spirits Scores
EPAS own scores	-0.57 $p < 0.01$	0.82 $p < 0.01$	-0.12 (N.S.)
EPAS predictions for villagers		-0.90 $p < 0.01$	0.22 $p < 0.05$
Alienation scores			-0.20 $p < 0.05$

All other correlations were statistically significant. The correlation of -0.57 between the student scores and their predictions is much stronger than the significant -0.26 obtained from the 1973 face to face data, and supports the finding in the original study that the higher on the Environmental Phenomena Attitude Scale a student scores, the more likely he is to depreciate village culture (Maddock, 1973).

The small but significant positive correlation between the Prediction scores and the Evil Spirit scores, and the samll but significant negative correlation between the Alienation and Evil Spirit scores indicates that despite the absence of any significant relationships between the Environmental Phenomena Scale and the Evil Spirit Scale, there is a small but distinct tendency for students low on the Evil Spirit Scale to depreciate their village culture; students with strong beliefs in evil spirits are slightly less likely to be alienated in terms of the attitude expressed by the Environmental Phenomena Scale.

The strong positive correlations between the Environmental Phenomena Scale and the Alienation scores and the very strong negative correlation between the Predictions and the Alienation scores are indicative of the degree of alienation present.

Analysis of variance effects. The results obtained from each of the four sets of scores were cast into a 2 x 2 x 2 analysis of variance design on regional (Highland – Island), educational (Forms 1 and 2 – Forms 3 and 4) and male – female dimensions, as indicated in Table 3.

In the case of the Evil Spirit Scale, no significant effects were found on the educational or sex dimensions, but an F ratio of 3.04 was obtained for the regional effect. This was significant at the 0.08 level, a slightly higher probability than the 0.05 level usually accepted in educational research. The means given in Table 4 show a substantial difference between East New Britain students and Southern Highland students at the Forms 1 and 2 level, with the East New Britain students being strong in their belief in evil spirits. The difference is quite small in Forms 3 and 4. The F of 3.04 produces an epsilon of 0.14 (Cohen, 1965, p. 104), which indicates that only a small percentage of the variance is associated with regional membership.

Forms 1 and 2 students in East New Britain are also higher on the Evil Spirit Scale than Forms 3 and 4, although the overall educational effect is not significant. The East New Britain result is in line with Stanton's findings in the 1969 study, but the Souther Highlands results are not.

TABLE 3

Analysis of Variance (2 x 2 x 2) design used, based on the following model:
Y = M + R + E + S + RE + SE + RS + ERS + ERROR

	R1 (Southern Highlands)		R2 (East New Britain)	
	S1 (Male)	S2 (Female)	S1 (Male)	S2 (Female)
E1(Forms) 3–4	N = 30	N = 8	N = 7	N = 8
E2(Forms) 1–2	N = 32	N = 7	N = 6	N = 5

TABLE 4

Mean Scores for Scales Used

*Region	Sex	Educational Level	Scale			
			Evil Spirit	EPAS Student	EPAS Predictions	Alienation
SH	M	Forms 1–2	25.9	63.6	45.0	18.6
		Forms 3–4	24.7	67.3	42.1	25.2
	F	Forms 1–2	27.8	63.3	46.3	17.0
		Forms 3–4	29.4	63.8	43.4	20.4
ENB	M	Forms 1–2	40.5	66.8	47.3	19.5
		Forms 3–4	27.5	68.0	47.6	20.4
	F	Forms 1–2	35.8	58.4	48.8	9.6
		Forms 3–4	31.9	65.9	47.5	18.4

* SH: Southern Highlands
ENB: East New Britain

In the case of the student scores on the Environmental Phenomena Scale, strongly significant educational ($F = 7.7$, $p < 0.01$) and sex ($F = 3.7$, $p < 0.005$) effects were obtained and the regional effect was not significant ($F = 0.06$). As was found in the group interviews in the original study (Maddock, 1973), Forms 3 and 4 students score significantly higher on the scale than Forms 1 and 2 students, with an epsilon of 0.25 (Cohen, 1965, p. 104) indicating that about six percent of the variance is associated with membership of educational level, and about eight percent associated with being male or female. These effects can be interpreted by looking at the means given in Table 4.

In the results obtained in the 1973 study males were consistently found to obtain higher mean scores, but the differences were not significant (Maddock, 1973).

In the case of the Prediction scores, the only significant effect obtained was the regional effect ($F = 3.9$, epsilon = 0.19, $p < 0.05$), there being no significant differences for educational level or sex. The mean values in Table 4 show that the Southern Highland students were predicting much lower scores for their minimally educated village counterparts than East New Britain students.

The only significant effect obtained for the Alienation score was in the education dimension ($F = 4.8$, $p < 0.05$, epsilon = 0.19) where Forms 3 and 4 students showed higher degree of alienation than Forms 1 and 2, as can be seen in Table 4.

Cultural alienation. When the students' own scores were compared with the predictions they made for answers expected of minimally educated villagers, a result very similar to that arrived at in the face to face interviews in the 1973 study (Maddock, 1973) was obtained. The difference of 20 scale points between the student mean and the prediction mean produced a t ratio of 18.7, significant at the 0.001 level. Mention has already been made of the significant negative correlation of -0.57 between the two sets of scores. In the 1973 study, the difference was 21.6, t 18.8, significant at 0.001, with a significant negative correlation of -0.26.

This is a clear indication of the level of cultural alienation, with respect to the attitude under study, of the high school student population in Papua New Guinea, a situation which has maintained its level over 3 years.

TABLE 5

Comparison of Student Scores with Predictions

	Mean	SD	Difference	t (dep)	p	r
Student Scores	65.1	5.1	20.0	18.7	<0.001	-0.57
Predictions	45.1	7.31				

Degree of dualism. The correlations between the scales show that there is no significant relationship between the attitude as measured by the Evil Spirit Scale, and that measured by the Environmental Phenomena Scale, although a small degree of relationship appears to exist between the Evil Spirit Scale scores and the degree of alienation. The analysis of variance indicates that in this sample educational level and sex are not significant factors influencing attitude to evil spirits. However educational level and sex are important factors for the Environmental Phenomena Attitude Scale, and educational level is a key factor in determining degree of alienation.

Some indication of the level of dualism can be obtained by casting the subjects into upper, middle and lower thirds, based on their scores on the two scales, as shown in Table 6. It can be seen from the table that about eight percent of the sample have significantly positive attitudes as measured by the Environmental Phenomena Attitude Scale, while professing strong beliefs in evil spirits as manifested by scores on the Evil Spirit Scale. Similarly about twelve percent show negative attitudes to investigation, control and manipulation of natural phenomena, but at the same

time have tended to choose natural explanations over an "evil spirit" explanation for the disappearance of the man in the Evil Spirit Scale story.

TABLE 6

Degree of dualism in attitude

	EPAS Upper Third	EPAS Middle Third	EPAS Lower Third	Total
Evil Spirits Upper Third	8 (7.8%)	13 (12.6%)	13 (12.6%)	34
Evil Spirits Middle Third	12 (11.6%)	11 (10.7%)	7 (6.8%)	30
Evil Spirits Lower Third	16 (15.5%)	11 (10.7%)	12 (11.6%)	39
TOTAL	36	35	32	103

Two items on the Environmental Phenomena Attitude Scale described situations where a person prone to belief in evil spirits could have interpreted them from a spirit orientation point of view. One item concerned the influence of a cave on gardening and other concerned a taboo being placed on people climbing a certain mountain because storms would come and destroy gardens as a result.

In the cave item, the negative attitude is reflected in choosing to abandon the land because of the cave. In the mountain item, the negative alternative indicates that the taboo would still be maintained, despite evidence that people climbed the mountain without ill effect. Table 7 summarizes the distribution of selection of the neutral and positive choices for each item between subjects in the upper, middle and lower thirds on the Evil Spirit Scale. A chi-square of 2.23 was obtained, not significant at the 0.05 level. This again indicates a degree of dualistic belief, consistent across both items.

TABLE 7

Selection of 2, 3 point choice by high, medium and low scorers on Evil Spirit Scale

EPAS	Upper Third	Middle Third	Lower Third
Cave Item	30	28	32
Mountain Item	34	31	30

χ^2 = 2.23 NS

Conclusions and implications

The Evil Spirit Scale, even in its present relatively unrefined form, reliably discriminates between individuals in the high school student population, along a continuum related to expressed choices of situations which compare a natural explanation with a "supernatural" explanation for the disappearance of a man in a feasible situation culturally similar to that which might occur in Papua New Guinea. The scores on the scale are interpreted as a measure of strength of attitude towards or strength of belief in the powers of undefined beings called "evil spirits".

The strength of belief, as measured by the scale, is strongly pervasive across the regions, the sexes and educational levels. The Stanton study of 1969 showed that Form 4 students decline in strength of belief over the lower forms, but in this study, the educational effect was not significant except that a distinct difference occurred in East New Britain.

The results obtained on the Environmental Phenomena Attitude Scale closely paralleled the results of the 1973 study, with Forms 3 and 4 students being significantly higher on the scale than Forms 1 and 2, except that the difference in favour of males was significant in this study, though a similar difference was not significant in 1973. The students maintained the trend found in 1973 to predict considerable difference in attitude between themselves and their minimally educated counterparts, this being interpreted as a measure of alienation. The alienation is stronger in the upper grades than in the lower, and there is a suggestion that the degree is greater in the Southern Highlands than in East New Britain.

No relationship was found between the scores on the Evil Spirit Scale and those on the Environmental Phenomena Attitude Scale, but a small tendency was detected for those scoring lower on the Evil Spirit Scale to score higher on the Alienation scale.

By looking at the performance of the upper, middle and lower thirds on each scale, distinct dualistic views were demonstrated, with about eight percent of the subjects holding "scientific" views on the environment, but strongly subscribing to belief in evil spirits, while a further twelve percent registered low on the Environmental Phenomena Scale but low on strength of belief in evil spirits.

It is not clear from this study whether the educational effect on views of the environment and, in the case of New Britain, on belief in evil spirits, is due to the effect of education itself, causing a shift in attitude, or whether the ones who survive selection at the end of Form 2 are the ones who hold such views in the first place. Further research is needed in this area. No specific claims can be made on the influence of science teaching.

However the Maddock (1973) study, through its open-ended follow-up interviewing, strongly indicated that science teaching was an important factor in changing peoples attitudes from more traditional to a "scientific" orientation. It is clear that a significant proportion of students at Form 4 level still subscribe to belief in evil spirits, despite the emphasis of the stated science syllabus aims concerning commitment to seek scientific explanation, and confidence in ability to affect changes and improvements in the environment. This dichotomy may be called dualism.

As stated earlier, failure to take account of traditional views and attitudes of the people in developing a formal school system can lead to tensions – there is clear evidence that alienation is occuring. This raises a major challenge for the future of curriculum development in Papua New Guinea.

This research is a relatively small scale, preliminary study, without rigid design controls, but its results confirm the findings of the Maddock (1973) study that cultural differences are being perpetrated by schooling. No work has been done using the Evil Spirit Scale on village people who

have not been to school. Even though the students score high on the scale, there could still be a degree of alienation between students and their village counterparts in this aspect of culture. Further research into this area could provide further cultural information of value to curriculum development.

References

COHEN, J. "Some statistical issues in psychological research." In Wolman, B.B. (Ed.), *Handbook of Clinical Psychology*. McGraw Hill, New York, 1965.

DART, F.E. & PRADHAN, P.L. "Cross cultural teaching of science." *Science*, **155**: 649-656, 1967.

DEPARTMENT OF EDUCATION PAPUA NEW GUINEA. *Secondary Science Syllabus, forms 1–4*, 1971.

EDWARDS, A.L. *Theory of Attitude Scale Construction*. Appleton Century Crofts, New York, 1957.

LAWRENCE, P. & MEGGITT, M.J. *Gods, Ghosts and Men in Melanesia*. Oxford University Press, Melbourne, 1965.

MADDOCK, M.N. "A pilot study to develop, refine and field test an instrument to measure the attitude of Papua New Guineans towards the investigation, control and manipulation of natural phenomena." Ph.D. Dissertation, Florida State University (USA), 1973.

MADDOCK, M.N. "Power over the environment – an attitude to the world", in *Educational Perspectives in Papua New Guinea*. Australian College of Education, 104-117, 1974(a).

MADDOCK, M.N. "Measuring attitudes held by high school students and minimally educated villagers in Papua New Guinea – an instrument development study." *Research in Science Education*, **4**: 113-129, 1974(b).

MADDOCK, M.N. "The attitude of Papua New Guineans towards investigation, control and manipulation of natural phenomena." *The Australian Science Teachers Journal*, **21** (*1*): 86-92, 1975(a).

MADDOCK, M.N. "The culture gap – what is formal schooling with its science education component doing to Papua New Guinea Society?" *The Australian Science Teachers Journal*, **21** (*1*): 93-97, 1975(b).

MALINOWSKI, B. *Magic, Science and Religion*. Beacon, Boston, 1948.

MEGGITT, M.J. "The Mae Enga of the Western Highlands", in Lawrence, P. & Meggitt, M.J., *Gods, Ghosts and Men in Melanesia*. Oxford University Press, Melbourne, 1965.

NUNNALLY, J.C. *Psychometric Theory*. McGraw Hill Series in Psychology, New York, 1967.

PRINCE, J.R. *Science concepts in a Pacific culture*. Angus & Robertson, Sydney, 1969.

STONE, R.H. "Science teaching in non-western societies." *The Science Teacher*, (United Kingdom), **10** (*5*): 21-23, 1967.

UNESCO. "Science and learning: The learning process and the teaching of science and mathematics in developing countries." *Report of Advisory Committee on the Application of Science and Technology to Development*. E/AC 52/3, 3 June, 1970.

VALENTINE, C.A. "The Lalakai of New Britain", in Lawrence, P. & Meggitt, M.J., *Gods, Ghosts and Men in Melanesia*. Oxford University Press, Melbourne, 1965.

USE OF LEARNING HIERARCHIES IN PROMOTING MASTERY LEARNING

Richard J. Trembath & Richard T. White

Most, if not all, senior secondary school physics courses include the skill of finding velocities at points on curves on position-time graphs. In Victoria this skill is part of the Grade 11 course. This paper describes an investigation into whether instruction based on a valid learning hierarchy can promote mastery of this skill by much younger children, in Grade 8. The investigation compares the achievement of the children learning the skill under different sequences of instruction and under different requirements of demonstration of mastery of pre-requisite skills.

Background

Mastery learning. Carroll's (1963) model of school learning lists certain requirements of the situation and of the learner before mastery can be achieved. These include:

1. A statement of the criteria which the learner must satisfy.
2. Sufficient time for learning.
3. The ability to understand the instruction.
4. A suitable quality of instruction.
5. Perseverance by the learner.

Bloom (1968) used Carroll's model to devise a practical system for mastery learning. Bloom argued that if students' aptitudes for learning are normally distributed, then to give each the same quantity and quality of instruction will ensure that their final achievements will also be normally distributed. Therefore in some way the instruction must take into account the attributes of the learner. While Bloom did much to popularise the idea of mastery learning, much remains to be done in discovering details of instructional procedures that will be effective in producing it. Learning hierarchies are perhaps the most promising lead there is to such procedures.

Learning hierarchies. Since the first postulation and study of learning hierarchies (Gagné, 1962), there have been several studies which support the basic premise that in order to learn certain skills one first must learn specific pre-requisite skills. Most of these studies are reviewed by White (1973).

A crucial point in the study of hierarchies was the distinction that Gagné (1968) made between verbalized knowledge and intellectual skills. A verbalized knowledge element is a single statement, such as knowing that second is a unit of time, and an intellectual skill is the capacity to perform a whole class of tasks, such as solving simple linear equations. Gagné suggested that learning hierarchies only apply to intellectual skills, which was later confirmed by White (1974).

White & Gagné (1974) argue that there is now enough evidence to support the validity of the learning hierarchy postulate for intellectual skills, and suggest that it is time to turn to other aspects of them, such as how they can be employed to promote mastery learning.

Learning hierarchies and mastery. Because the development and validation of a learning hierarchy enforces precise behavioural definition of the terminal skill and produces a detailed description of the pre-requisite skills, a hierarchy meets, or makes it simple to meet, several of Carroll's requirements.

The skill which the student has to learn is narrowly defined, so all that remains to satisfy

the first requirement is to specify how many times he has to demonstrate it before he is accepted as having learned it.

A hierarchy makes it probably that the second requirement, of sufficient time for learning, will be met, because it contains no irrelevant skills, and therefore is of maximum efficiency, and because the sequence of instruction is clear and so prevents loss of time while students batter away at impossible tasks for which they lack the essential pre-requisites.

Students should be able to understand the instruction associated with a hierarchy, because at every point it builds on what they already can do and moves them on a small step further.

The fourth requirement, a suitable quality of instruction, can be met by giving students help with each skill, and ensuring that they do not proceed to higher, and consequently impossible tasks, until they have demonstrated their mastery of the one at hand. Instruction based on a hierarchy can readily have diagnostic and achievement tests for individual skills as an integral part.

Hierarchy theory has nothing to say about perseverance, Carroll's fifth requirement, except that the careful sequencing of skills and the appropriate size of the learning step in going from one skill up to another ensures much success, which should encourage perseverance. Although there is no objective evidence to support this statement, it is the authors' observation that students working on a learning program which is based on a hierarchy show surprisingly high concentration on the task.

Previous experiments on mastery learning. There have already been some studies in which a learning hierarchy was used to plan the course of a mastery learning sequence. However, they suffer from one or more of the following weaknesses:

1. The hierarchy was not properly validated.
2. The elements in the hierarchy were so loosely defined that they could have contained more than one skill.
3. No distinction was made between verbalized knowledge and intellectual skills.

An example of how hierarchies are obtained for these studies is provided by Airasian (1970, 1971). He had two experienced teachers independently postulate a learning hierarchy for the material in a chemistry textbook, and found a 90% agreement between elements in their hierarchies. As the "teachers were cautioned to describe only what was present in the chapters under study" (1971, p. 35), a high level of agreement could be expected. If the teachers had been free to choose the subordinate elements required to achieve a certain criterion Airasian might have found less agreement. The subordinate elements in the hierarchy were large and ill-defined (e.g. "coulomb" and "wavelength"), which can hide differences in understandings of meanings of the elements.

Most of the research on mastery learning has involved untested assumptions about the subordinate skills, and their sequence of presentation, which a learner must have in order to succeed at a task. Twenty-seven of the mastery learning studies reviewed by Block (1971) could have made use of a validated hierarchy, but in only one case was there an attempt to do so.

Sequence of instruction. The general result of studies of sequence of instruction, in which a forward or logical sequence is compared with random, or scrambled, and reverse sequence, has been one of no significant difference. This result has been a puzzle for some time, since it appears contrary to what might reasonably be expected. The explanation may be found in methodological flaws in the experiments. The studies of Roe, Case, & Roe (1962), Levin & Baker (1963),

Hamilton (1964) and Payne, Krathwohl & Gordon (1967) suffer from one or more defects: no distinction was made between verbalized knowledge and intellectual skills; the forward sequence was not based on a valid hierarchy; or the subject matter was either too easy or too difficult for the learners.

More recently, Brown (1970) found that a logical sequence group performed better than scrambled and reverse sequence groups on a number series task which involved intellectual skills. Resnick, Siegel & Kresh (1971) and Caruso & Resnick (1972) found that kindergarten children who followed a hierarchical sequence learned classification skills in fewer trials than children who followed different sequences. This result was used as evidence for the validity of the hierarchy, however, rather than as evidence that a valid hierarchy represents best sequences.

Demonstration of mastery of each element. Besides varying the sequences of presentation of elements to be learned, experiments may require learners to demonstrate acquisition of each element as it is presented, or may allow them to move on rapidly without check to the end. The hierarchy model leads to a prediction that if learners are required to demonstrate mastery of each element before being permitted to advance to the next one, they will be better able to learn the next element. Merrill, Barton, & Wood (1970), working with a logically derived though not empirically validated hierarchy, found that learners given a specific review treatment for each element that they failed to learn at first attempt worked through the entire program in less time and took less time to complete a transfer test than learners who were simply given a right-wrong feedback for each task. There is, however, little other evidence that bears on the prediction. There is a need for a direct test of whether a stringent requirement of mastery of each element in a hierarchy before progression enhances students' performance over the entire hierarchy.

The present study

The study examined the effects of sequence of instruction and mastery of subordinate elements on the learning performance of Grade 8 students on a hierarchy of graphical skills validated by White (1971) (see figure 1). Perhaps more important than the questions involving the above two independent variables is the question of whether there is a chance of achieving 100% mastery of the elements by 100% of the students in Grade 8. Such a question does not fit the standard models of educational research, but is important to answer since it is at the heart of the purpose of educational practice. White & Gagné (1974) discuss the importance of the question and ways in which attempts may be made to answer it. In essence their recommendation is to choose a task for which there are *a priori* reasons for judging it a difficult one for the chosen group of learners, and to show that the instructional treatment produces unexpected success. In this experiment the *a priori* reasons are that normally the subject matter of the hierarchy is taught to a select group three years older than the children in the experiment, and that diagnostic test norms for the select group, and their performance on these skills in external examinations, show that they have quite imperfect learning of the skills.

The eighth grade students in the experiment came from six outer suburban high schools in areas of Melbourne that are suggested by Lancaster Jones (1969) as being near the mean for socio-economic factors. One class in each school was selected and randomly divided into four treatment groups – two levels of sequence, forward and reverse; and two levels of demonstration of mastery of subordinate skills.

Forward groups worked through the skills in the order XII XI X VIII VII VI III XXIV XXIII XXII XXXI XXIX XXVIII XVII XIV I. Reverse groups worked through

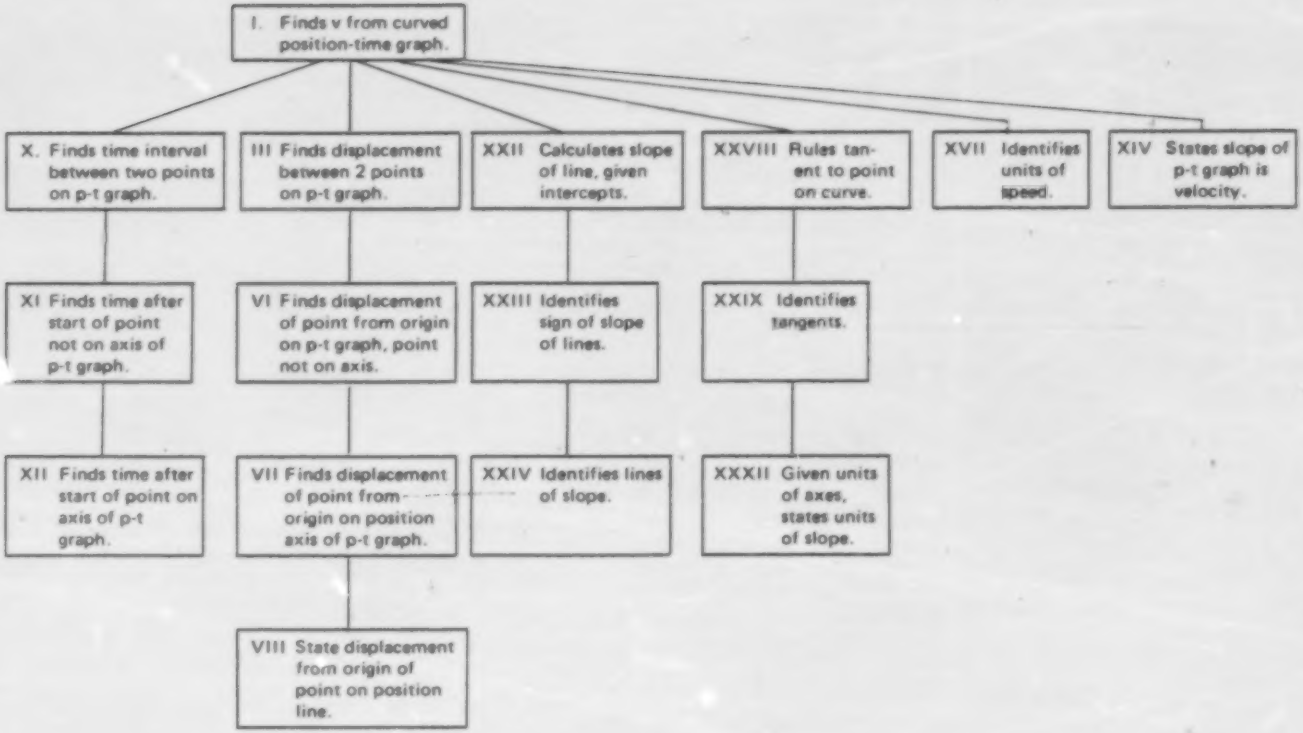

I. Finds v from curved position-time graph.
X. Finds time interval between two points on p-t graph.
III Finds displacement between 2 points on p-t graph.
XXII Calculates slope of line, given intercepts.
XXVIII Rules tangent to point on curve.
XVII Identifies units of speed.
XIV States slope of p-t graph is velocity.
XI Finds time after start of point not on axis of p-t graph.
VI Finds displacement of point from origin on p-t graph, point not on axis.
XXIII Identifies sign of slope of lines.
XXIX Identifies tangents.
XII Finds time after start of point on axis of p-t graph.
VII Finds displacement of point from origin on position axis of p-t graph.
XXIV Identifies lines of slope.
XXXII Given units of axes, states units of slope.
VIII State displacement from origin of point on position line.

FIGURE 1: Hierarchy of Graphical Skills

them in the order I X XI XII III VI VII VIII XXII XXIII XXIV XXXI XXVIII XXIX XVII XIV and I again.

One of two levels of demonstration of performance on each subordinate element was required of each learner. For each element he had to obtain either one correct answer or three consecutive correct answers before proceeding.

The instruction sheet for each element was stapled to a large number of items testing mastery of that element, and then bundles of papers were given to the students in appropriate sequences. Reverse sequence bundles were collected in folders with the highest element at the top. Forward sequence learners worked straight through their bundles. The reverse sequence learners worked down their bundles until they reached mastery on an element, when they worked up to the top of that particular column again. For example, a reverse sequence student might fail to reach mastery immediately on elements III and VI, but reach it on element VII. He would then ignore element VIII, and return to element VI, attain mastery of it, then return to element III to attain mastery of it.

To compare performances, the grade 11 Physics group in each of the six schools was given the same set of pre-and post-tests. The form of this comparison is illustrated by Table 1.

TABLE 1

Comparison of testing and learning procedures for Grade 8 and 11 students

	Test group – Grade 8	Comparison group – Grade 11
Pre-test	Co-operative reading test. Maths skills. Final skill of hierarchy.	As for Grade 8
Instruction	Learning program Sample split into 4 groups. F1 F3 R1 R3.	By class teacher.
Post-test	3 items on final skill. Administered on completion of learning program.	3 items administered in conjunction with normal test.

Results and discussion

Performance of Grade 8 learners on learning program. Learners were required to show one of two levels of mastery on element 1 of the learning program. F1 and R1 learners had to give the correct answer to 1 item on element 1 while F3 and R3 learners had to give the correct answer to 3 consecutive items on element 1. The results are shown in Table 2.

No comparison of F1 and F3 groups can be made as a different criterion of achievement was used in each case. Similarly no comparison of R1 and R3 groups can be made. It could be argued that F1 (33/36 reached criterion) is marginally superior to R1 (32/38) and that F3 (35/38)

is superior to R3 (27/37) i.e. the forward sequence is slightly better in producing mastery than the reverse sequence. Of much greater significance is the fact that 65/74 (88%) learners achieved the 1 item correct level of mastery and 64/75 (85%) learners achieved the 3 item level of mastery, on a task which even Grade 11 learners, taught by usual methods, find difficult (see below).

TABLE 2

Performance of Grade 8 learners on elemenb 1 of learning program

	F1	F3	R1	R3
Numbers achieving criterion	33	35	32	27
Numbers not achieving criterion	3	3	6	10

As learners worked through the program assistance was given to them when requested. When giving assistance the instructor introduced no new information to the learner but simply directed his attention back to the appropriate instruction sheet. After receiving assistance the learner was required to get 1 or 3 more items correct.

There is always the possibility that learners cheated by simply copying the answers from the self-correcting program. On element 1 this was prevented by the instructor observing the calculations before allowing the learner to take the post-test. A post-test gives reliable evidence of learner performance.

Performance of Grade 8 and Grade 11 learners on post-test. Both Grade 8 and Grade 11 learners were given the same post-test of 3 items on element 1.

Although the criterion of mastery was to give 3 correct answers the total number of correct answers of all learners was recorded also (see Table 3).

TABLE 3

Performance of Grade 8 and Grade 11 learners on post-test

	Grade 8				Grade 11
	F1	F3	R1	R3	
Number of learners	36	38	38	38	101
Number achieving mastery	10	16	6	12	47
Number not achieving mastery	26	22	32	25	54
Number of items correct (includes masterers and and non-masterers)	59	75	35	54	206

On this test 16/38 (42%) of the F3 and Grade 8 learners were able to achieve the defined level of mastery yet only 47/101 (4.. %) of Grade 11 learners achieved the same level of mastery on the same test when included in the normal test program for their Physics class.

When judged on number of items correct the performance of the F3 group was 75/114 (66%) which compared favourably with the Grade 11 group 206/303 (68%).

The performance of F1 and R3 is about the same when judged both by numbers achieving mastery and numbers of items correct.

Presumably the more rigorous criterion of mastery within the R3 learning program offset the disadvantage of reverse sequence. The F3 group had a superior performance on the post-test than the F1 and R3 groups and the R1 group had an inferior performance to the F1 and R3 groups.

It should be noted that even the R1 and R3 groups were presented with all subordinate elements and learners were able to deduce the forward sequence and follow its plan of instruction if they needed to do so.

Conclusion

Although it is interesting that the F3 group performed better than the F1 and R3 groups and that these two groups performed better than the R1 group this is hardly surprising.

What does seem to be important is the performance of the Grade 8 learners when compared with the Grade 11 learners.

Although this learning task is a difficult one for Grade 11 learners the Grade 8 learners performed almost as well when judged on numbers achieving mastery and on numbers of items correct. Even more interesting is the high percentage of learners who achieved a defined level of mastery on element 1 of the learning program.

References

AIRASIAN, P.W. "The use of hierarchies in the analysis and planning of chemistry instruction." *Science Education*, **54**: 91-95, 1970.

AIRASIAN, P.W. "The use of hierarchies in cirriculum analysis and instructional planning." *California Journal of Educational Research*, **22**: 34-41, 1971.

BLOCK, J.H. (ed.). *Mastery Learning: Theory and Practice*. Holt, Rinehart and Winston, New York, 1971.

BLOOM, B.S. "Learning for mastery." *Evaluation Comment*, **1**, 1968.

BROWN, J.L. "Effects of logical and scrambled sequences in mathematical materials on learning with programed instruction materials." *Journal of Educational Psychology*, **61**: 41-45, 1970.

CARROLL, J.B. "A model of school learning." *Teachers' College Record*, **64**: 723-733, 1963.

CARUSO, J.L. & RESNICK, L.B. "Task structure and transfer in children's learning of double classification skills." *Child Development*, **43**: 1297-1308, 1972.

GAGNÉ, R.M. "The acquisition of knowledge." *Psychological Review*, **69**: 355-365, 1962.

GAGNÉ, R.M. "Learning hierarchies." *Educational Psychologist*, **6**: 1-9, 1968.

HAMILTON, N.R. "Effects of logical versus random sequencing of items in an auto-instructional program under two conditions of covert response." *Journal of Educational Psychology*, **55**: 258-266, 1964.

LANCASTER JONES, F. *Dimensions of Urban Social Structure*. A.N.U. Press, Canberra, 1969.

LEVIN, G.R. & BAKER, B.L. "Item scrambling in a self-instructional program." *Journal of Educational Psychology*, **54**: 138-143, 1963.

MERRILL, M.D., BARTON, K. & WOOD, L.E. "Specific review in learning a hierarchical imaginary science." *Journal of Educational Psychology*, **61**: 102-109, 1970.

PAYNE, D.A., KRATHWOHL, D.R. & GORDON, I. "The effect of sequence on programmed instruction." *American Educational Research Journal*, **4**: 125-132, 1967.

RESNICK, L.B., SEIGEL, A.W. & KRESH, E. "Transfer and sequence in learning double classification skills." *Journal of Experimental Child Psychology*, **11**: 139-149, 1971.

ROE, K.V., CASE, H. & ROE, A. "Scrambled versus ordered sequence in auto-instructional programs." *Journal of Educational Psychology*, **53**: 101-104, 1962.

WHITE, R.T. "Learning graphical skills in kinematics." Ph.D. thesis, Monash University, 1971.

WHITE, R.T. "Research into learning hierarchies." *Review of Educational Research*, **43**: 361-375, 1973.

WHITE, R.T. "The validation of a learning hierarchy." *American Educational Research Journal*, **11**: 121-136, 1974.

WHITE, R.T. & GAGNÉ, R.M. "Past and future research on learning hierarchies." *Educational Psychologist*, **11**: 19-28, 1974.

THEMATIC STRUCTURAL ANALYSIS APPLIED TO CONCEPTUAL HIERARCHIES

Brian W. Carss

The purpose of this paper is to describe and illustrate a procedure for extracting a graphical representation of a hierarchical arrangement of elements from a set of elements. The procedure is an extension of the work by Carss (1973) on Thematic Structural Analysis which enables an author or curriculum developer to establish a priority among the concepts that are represented by thematic elements.

In studying complex problems, such as those normally encountered in educational research, it is desirable, and at times essential, to attempt to build a hierarchy. People have recognized this fact, and usually attempt to build the hierarchy intuitively, thereby constraining themselves unnecessarily. In order to cope with the complexity in assembling a hierarchy, important elements are eliminated as a simplifying gesture and such simplification may destroy the value of the hierarchical arrangement.

There are three concepts of a hierarchy that require definition in the present setting; the first consists of elements that are to be arranged into a hierarchy, the second is the content or relationship between any pair of elements, and the third is direction. Since any hierarchical arrangement is directional, it can be regarded as a subordination relationship. Elements lower in the hierarchy can be regarded as being subordinate to elemen located higher.

The major task in developing any hierarchy is to identify the relevant set of elements and having done this, then to determine their subordinate relationships and encode them. In this study, however, there is not any difficulty in determining the element set because the thematic elements extracted by the thematic structural analysis technique (TSA) (Carss 1973) form the element set.

The problems associated with determining subordination relationship are also alleviated as these are automatically determined by the TSA technique. What remains then, is the encoding of the data, and the extraction of the hierarchy, and what follows is a description of the methodology to accomplish these latter two tasks.

Suppose that the thematic elements have been numbered 1 to n, then a square, n x n matrix is constructed as shown in Figure 1.

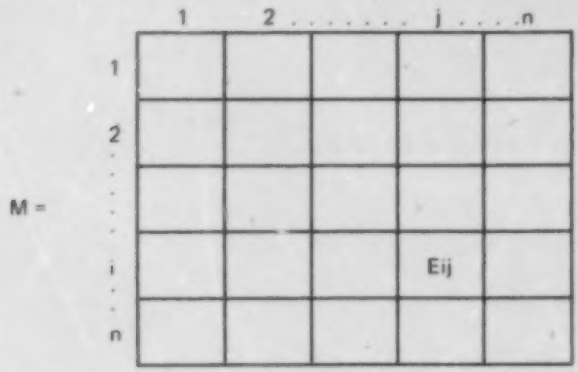


FIGURE 1

The entry in the i th row and the j th column will reflect the subordination relationship between the two thematic elements i and j. If thematic element i is subordinate to j, then a 1 will be recorded in **Eij**, and if i is *not* subordinate to j, a zero is recorded in the matrix.

In addition, the construction of a subordination matrix is governed by a set of three rules which are as follows:

(1) All entries on the main diagonal (**Eii**) are equal to zero.

(2) If element i is subordinate to element j, a 1 is entered in position **Eij** and it follows that a zero must be entered in cell **Eji**.

(3) If element i is subordinate to element j, and element j is subordinate to element k, then element i is subordinate to element k. This is known as the transitivity condition.

Illustrative example. This example is taken from Warfield (1973) to illustrate how a hierarchy is extracted from a simple system of four elements. The four elements consist of the following objectives:

1. To lose weight; 2. To diet; 3. To exercise daily; and 4. To swim daily.

The subordination matrix will therefore be a 4 x 4 one.

Applying rule (1), the diagonal is filled with zeros (Figure 2.1).

	1	2	3	4
1	0			
2		0		
3			0	
4				0

FIGURE 2.1

	1	2	3	4
1	0			0
2		0		
3			0	
4	1			0

FIGURE 2.2

A daily swim would contribute to a loss of weight, therefore $E_{41} = 1$, and applying rule (2), $E_{14} = 0$, (Figure 2.2).

	1	2	3	4
1	0			0
2		0		
3			0	0
4	1	0	1	0

FIGURE 2.3

	1	2	3	4
1	0	0	0	0
2	1	0	0	0
3	1	0	0	0
4	1	0	1	0

FIGURE 2.4

A daily swim would not be subordinate to dieting, so objective 4 is not subordinate to

objective 2. The entry in E_{42} must be zero; nothing is implied about any of the other entries however (Figure 2.3). A daily swim is subordinate to daily exercise, thus $E_{43} = 1$ and $E_{34} = 0$ (Figure 2.3). This procedure is continued until all of the cells in the matrix are filled (Figure 2.4).

The example matrix has four elements, so the hierarchy will also contain the same number of elements. Wherever there is a 1 in the matrix there will be a connection between those two elements. Applying this procedure, one can see that the matrix in Figure 2.4 can be resolved in the hierarchy Figure 3.1.

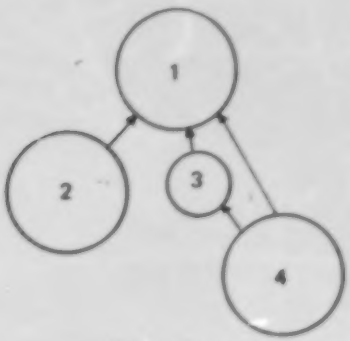


FIGURE 3.1

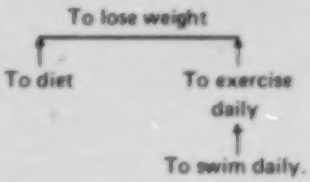


FIGURE 3.2

However, the diagrammatic representation of the hierarchy as shown in Figure 3.1 is not satisfactory as it contains redundant information. There is no need for the connection between 4 and 1 because 4 and 1 are connected through objective 3. Figure 3.2 is simpler and all of the important subordination conditions have been retained.

Thematic elements and their hierarchy

As the first test of the procedure outlined above with elements extracted by the Thematic Structural Analysis technique, chapter 3 from the book *Spatial Organization* by Abler *et al.* (1971) was selected. The chapter was entitled "The Science of Geography" and in it, the viewpoint that geography is a science is supported and developed. The concept of location is fundamental to all geography and it is directly related to the concept of distribution and the way in which distributions relate to each other through covariation and interaction. The modern approach to geographical study involves explanation of human spatial behaviour as the product of the relative spaces which man himself creates by his space-adjusting activities.

Twenty three thematic elements were extracted from this chapter (Table 1); six referred to location; thre to distribution; two each to region and change through time, and four to spatial interaction. The remaining six were classified as non-geographical themes.

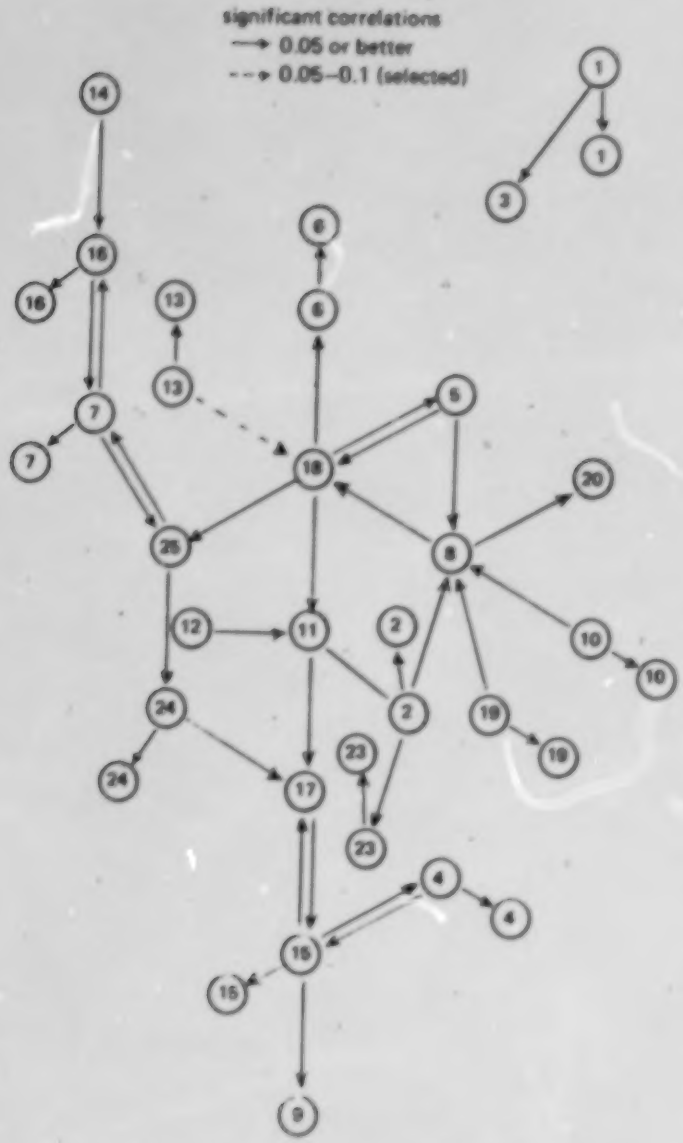
Examination of the idealized structural diagram shows a complex pattern of cross-correlations between the thematic elements, indicating that they are closely related to one another. Some themes cross-correlate with a number of other themes, indicating the high degree of internal interdependance of the ideas. There is a progression of themes beginning with the concept of location, and then progressing through the concept of region, spatial association and interaction.

Figure 5 is the hierarchy of thematic elements as extracted by the procedure described earlier. In general, the simplest concept of location is located at the bottom of the hierarchy while,

TABLE 1

Thematic Elements – Abler, Adams, Gould

THEMATIC ELEMENTS	GEOGRAPHICAL CONCEPTS
1. The development of geography and cartography in the early Christian world, the medieval Muslim period, and in the Far East.	Change through time
2. Various sciences may be distinguished by their focus on different questions about the world.	
3. Geographers ask and attempt to answer 'Where?' questions about phenomena.	Location
4. Geographical concepts (such as distance) may be considered either in absolute or relative contexts; for example, in relation to cost of distance.	Location
5. A major concern in geography is with distributions observable from local to terrestrial scales.	Distribution and scale
6. Spatial structure and spatial process interact causally.	Spatial interaction
7. Uniform and nodal regions.	Region
8. Spatial distributions may occur in one or multi-dimensional spaces.	Distribution
9. Human spatial behaviour.	Spatial interaction
10. Problems of organising new knowledge (about locations) in geography.	Location
11. Systems of thought – problems and methods – in geographical study.	
12. Geographers are concerned with accurate measurements.	
13. Ways of describing location.	Location
14. Geographers are interested in What is where? questions.	Location
15. Ways in which geographers use information.	
16. Data gathering for answers to What is where? questions	Location
17. Change in the character and use of maps relate to the development of the main concerns of geographers.	Change through time
18. Geographical analysis is concerned with the distribution of phenomena in terrestrial space.	Distribution
19. Changes in relative space influence human spatial behaviour.	Spatial interaction
20. Travel-time and relative distance.	Spatial interaction
23. Constructs may relate to the theoretical structures of several sciences.	
24. Modern geography (1800-1950) emphasised the delimitation of regions of like phenomena.	Region
25. Geographical thought about the earth's surface.	



Numbers refer to thematic elements

FIGURE 4: Abler *et al.* Thematic Structure Idealised

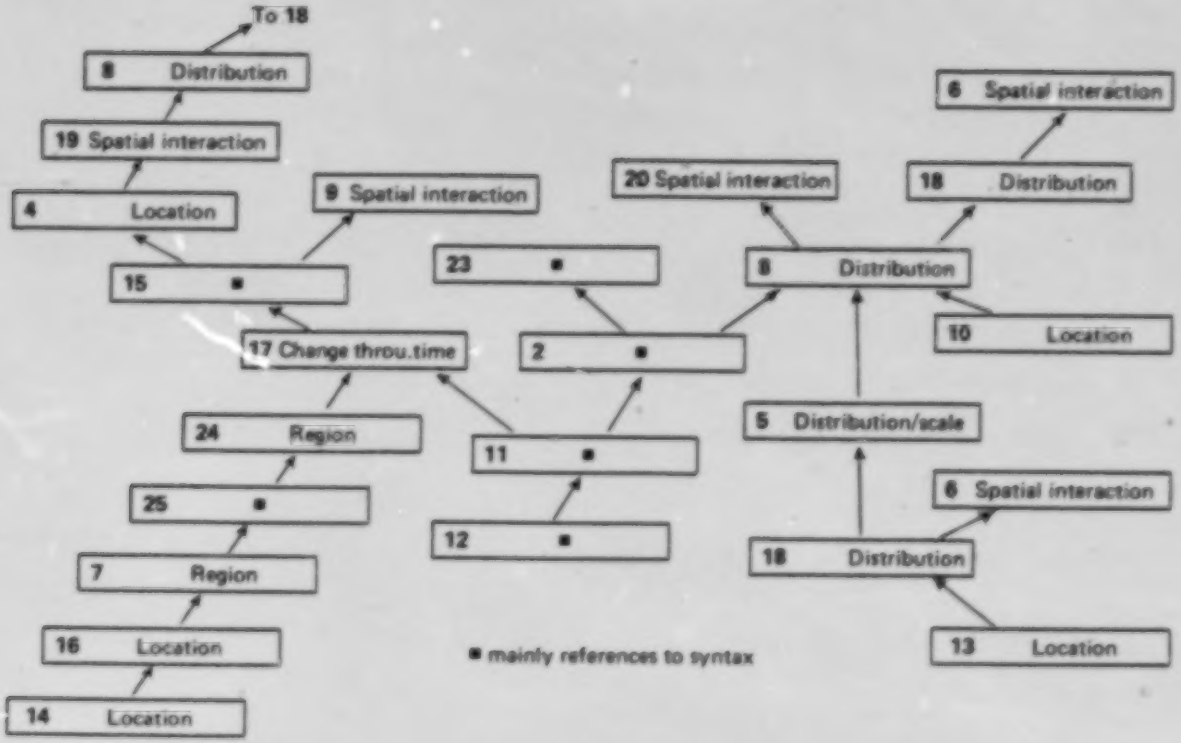


FIGURE 5: Hierarchy of concepts (idealised) Abler *et al.*

the more complex concept of spatial interaction occurs at the top. Another interesting feature of this diagram is the central position for the thematic elements concerned with syntax in geography, emphasizing the relations between concepts, questions and inquiry.

A similar analysis was carried out on Chapter 1 of *Educational Psychology: A Cognitive View* by David P. Ausubel, 1968 (Table 2). The analysis indicated that there were four structural characteristics which become more apparent as the chapter developed. The first of these characteristics is that certain themes are given extended discussion. These themes consist of specialized ideas which contribute to the detail in the substance of the chapter.

TABLE 2

Thematic elements: Ausubel, Chapter 1

1. Factors affecting the development of intellectual ability, for example, readiness, motivation, social class background.
2. Types of learning: meaningful and rote, reception and discovery.
3. The school's concern for the intellectual development of students following an academic curriculum.
4. Research orientations – basic science versus applied science approaches – and their relevance to educational problems.
5. Relations between commonsense and theory in educational psychology.
6. Functions of discovery learning in relation to the solution of problems and the acquisition of subject matter knowledge.
7. Developmental needs of children.
8. Relationships between knowledge, and methods of learning, teaching and researching.
9. Relationships between theories of learning, theories of teaching and educational psychology.
10. Teaching subject matter to pupils, including reference to the effects of motivation.
11. Verbal learning and the processes affecting the acquisition of concepts.
12. Laws relating to classroom learning.
14. Acquisition of knowledge by discovery; including the implications for pedagogy.
15. Instructional conditions and variables.
17. Organisation (for example, of subject matter) for teaching, learning and researching.
18. The principle of readiness in relation to teaching methods and curriculum.
19. Experimental findings about student learning.
20. Cognitive and social factors as variables affecting learning.
21. Values of different approaches to learning and to research.
23. The generality of laws derived from research findings into variables affecting learning.
24. The effectiveness and utility of knowledge about variables affecting educational practices.

The second characteristic relates to the first, in that some themes are not introduced until the chapter is well advanced, and the third is concerned with the degree of cross-correlation between the themes, and how the webs formed by the cross-correlating elements becomes more complex as the chapter progresses. The intricate pattern of themes emphasizes the complexity of the substantive structure and also the elaboration of ideas. This is consistent with Ausubel's principle of progressive differentiation.

The fourth characteristic concerns the use of advance organisers where themes are introduced

early in the chapter and then elaborated upon later. One can conclude that the author is faithful to the structural principles outlined in the foreword of his book. There are clear indications of progressive differentiation, of the use of advance organizers and of integrative reconciliation.

The hierarchical diagram of these thematic elements (Figure 6) illustrates very well the structural characteristic of progressive differentiation. The thematic elements occurring at the bottom of the diagram are all very general and encompassing concepts. For example, thematic element 8 relates knowledge, methods of learning, teaching and research. Similarly, thematic element 9 is concerned with the relationship between theories of learning, teaching and educational psychology. However, by the time one reaches the top of the hierarchy, very specific thematic elements appear. Thematic elements such as 11, relating verbal learning to the acquisition of concepts; 7, which discusses the developmental needs of children and 19, which presents experimental findings about student learning.

What is also of interest in this diagram are the thematic elements that have not been included. There are two trivial hierarchies, one containing two elements 15 and 14 and the other only one element 6. These three elements are all concerned with discovery learning, which, from the author's viewpoint, are not consistent with the progressive differentiation of ideas which is so very obvious in Figure 6.

The existence of trivial hierarchies suggests a lack of coherence in the writing between the main themes and those subordinate themes which appear to be trivial.

In conclusion, the methodology that has been described works well, and the resulting hierarchies appear to be consistent with the stated aims of the authors. There is obviously a need to develop the procedure further in order to handle non-subordinate or bi-directional relationships. This proposed extension would lead logically to new techniques for interaction analysis and cluster analysis.

Acknowledgements

I would like to thank Mr J. Clarke and Mr G.B. Cox for their help, and for the permission to use some of their materials in the preparation of this paper.

References

ABLER, R., ADAMS, J.S. & GOULD, P. *Spatial Organizations: The Geographer's View of the World.* Prentice Hall, Englewood Cliffs, 1971.

AUSUBEL, D.P. *Educational Psychology: A Cognitive View.* Holt, Rinehart and Winston, New York, 1968.

CARSS, B.W. "Content analysis as a technique for determining the structure of materials." *Science Education: Research 1973,* 107-117, 1973.

WARFIELD, J.N. "On arranging elements of a hierarchy in graphic form." *IEEE Transactions on Systems, Man and Cybernetics,* 3: 121-132, 1973.

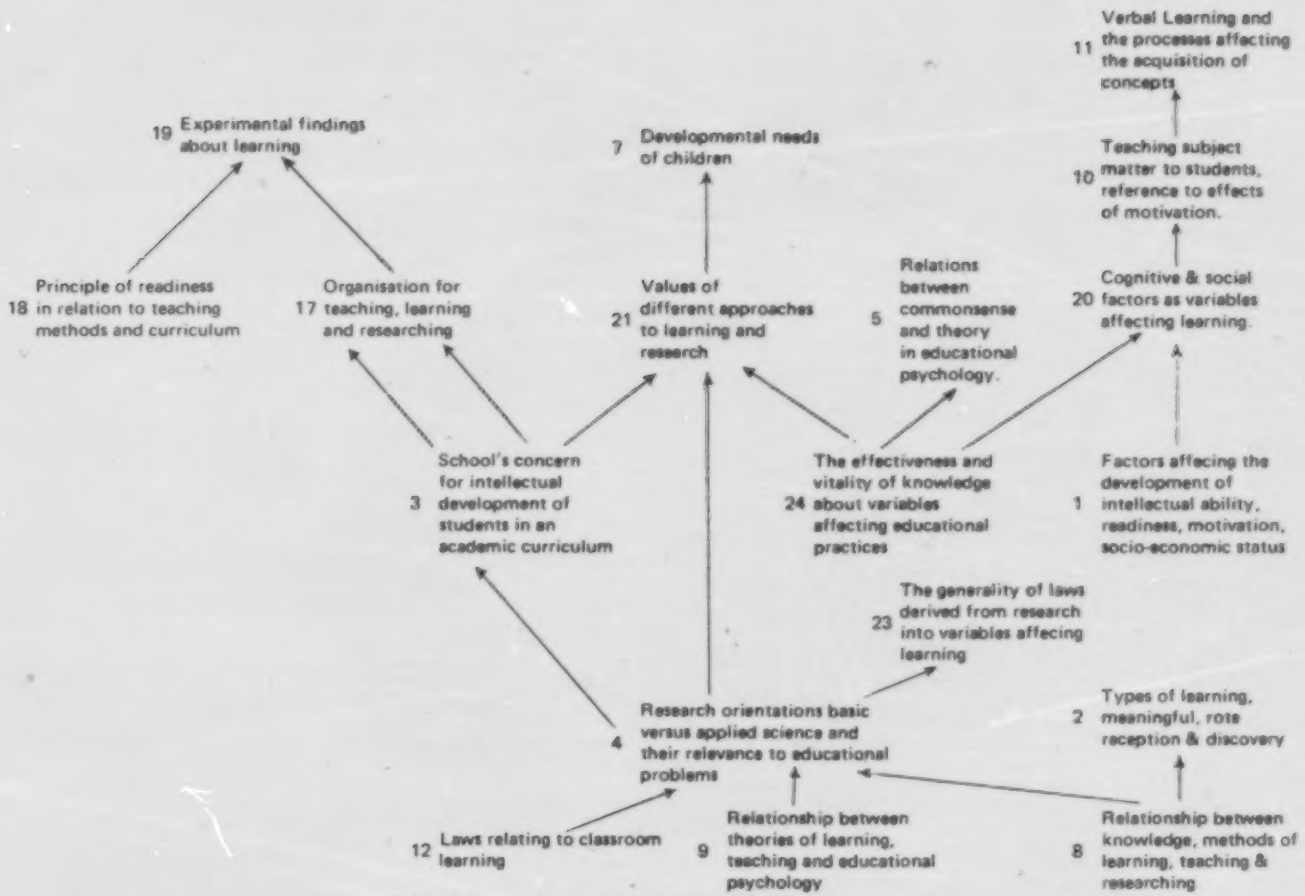

11 Verbal Learning and the processes affecting the acquisition of concepts
19 Experimental findings about learning
7 Developmental needs of children
10 Teaching subject matter to students, reference to effects of motivation.
Principle of readiness
18 in relation to teaching methods and curriculum
Organisation for
17 teaching, learning and researching
Values of
21 different approaches to learning and research
5 Relations between commonsense and theory in educational psychology.
Cognitive & social
20 factors as variables affecting learning.
School's concern for intellectual
3 development of students in an academic curriculum
The effectiveness and
24 vitality of knowledge about variables affecting educational practices
1 Factors affecting the development of intellectual ability, readiness, motivation, socio-economic status
23 The generality of laws derived from research into variables affecting learning
Research orientations basic
4 versus applied science and their relevance to educational problems
2 Types of learning, meaningful, rote reception & discovery
12 Laws relating to classroom learning
9 Relationship between theories of learning, teaching and educational psychology
8 Relationship between knowledge, methods of learning, teaching & researching

FIGURE 6: Hierarchy of Thematic Elements in Ausubel (Chapter 1, "Educational Psychology: A Cognitive View")

CAVEAT MENTOR – LET THE TEACHER BEWARE!

E. Jungwirth

> A cautious man should above all be on guard against resemblances – they are a very slippery sort of thing.
>
> (The stranger from Elea in Plato's *Sophist*)

The use of analogy as a teaching-device is extremely common in science-teaching. It is a way of "reducing" the "unfamiliar" to the familiar. However, this way out of a difficulty can, and does, frequently boomerang, quite apart from the fact that such reduction cannot be regarded as explanation. In this paper several quotes from the literature on scientific explanation will be cited in order to bring "analogy-explanations" into proper perspective. Following this, some research results, relating to one particular type of analogy, anthropomorphic formulations, will be reported, and finally some recommendations will be made, which may obviate some of the dangers of mistakenly accepting analogies for identities.

Examples from the literature

Weaver (1964) distinguishes between two main kinds of explanations. The first type, consisting of explaining something by restating or describing the unfamiliar by the analogy employed, is examplified by the "explanation" of the spreading of electromagnetic waves by the analogy of the results of dropping a pebble into water. This he regards as a reasonable example of an explanation many people would find intellectually satisfactory. Braithwaite (1960) would call this a "first-stage" explanation. Bronowski (1959) states: "The scientist or the artist takes two facts or experiences which are separate; he finds in them a likeness, which had not been seen before; and he creates a unity by showing the likeness". This idea is, indeed, very much older, but just as old are the warnings against the danger involved. Plato (see above) and Francis Bacon: ("The human intellect from its peculiar nature, easily supposes a greater uniformity and equality in things than it really finds . . . . it feigns parallels, correspondents and relations which do not exist" (*Novum Organum*, Book 1)), have warned against uncritical acceptance of analogies, in spite of the "pleasure which the mind derives from the perception of similitude in dissimilitude" (Wordsworth – *Lyrical Ballads*).

Hempel and Oppenheim (1948) analyse the "familiarity" type of explanation whilst referring to teleological explanations:

> A teleological explanation tends to make us feel that we really understand the phenomenon in question, because it is accounted for in terms of purposes, with which we are familiar. But it is important to distinguish here understanding in the psychological sense of a feeling of empathic familiarity, from understanding in the theoretical, or cognitive, sense of exhibiting the phenomenon to be explained as a special case of some general regularity (this refers to the Hempelian "cover-law" nomological explanation-rules). The frequent insistence, that explanation means the reduction of something unfamiliar to ideas already familiar to us is indeed misleading. For while some scientific explanations do have this psychological effect, it is by no means universal: The free fall of a physical body may well be a more familiar phenomenon, than the law of gravitation, by means of which it is explained (again – nomological explanation by

> subsumptive law); and surely the basic ideas of the theory of relativity will appear to many to be less familiar, than the phenomena for which the theory accounts. "Familiarity" of the explicans is thus not only not necessary for a sound explanation, but it is not sufficient either. This is shown by the many cases in which a proposed explicans sounds suggestively familiar, but on closer inspection proves to be a mere metaphor . . . .

Hempel and Oppenheim are of the opinion that it is their anthropomorphic character which lends the appeal to teleological explanations. That teleological explanations ("in order to", "for the sake of" etc.) are *intrinsically* anthropomorphic is hotly disputed in the philosophical literature, but a discussion of this topic would be beyond the scope of this paper. However, the philosophical aspect of this problem is probably secondary – at least in high-school teaching – to another problem emanating from Hempel and Oppenheim, but not followed up there. Do pupils (and teachers?) recognize a metaphor when they see one? In other words, do they recognize an analogy for what it is, or do they accept it, mistakenly, for an identity.

Several biologists and philosophers of science have, while discussing "weightier" problems, remarked on the importance of being able to analyse statements critically and not to be misled by offered analogies. Nagel (1961) states:

> When familiar notions are extended to novel subject matter on the basis of unanalysed similarities, serious error can easily be committed. Animistic explanations of physical events are well-known illustrations of such unwarranted extensions of conceptions from a domain in which their use is legitimate, to domains in which it is not.

Ghiselin (1967) says: "Such notions as the apparent purposefulness are mere pseudo-problems to those fully aware of the dangers of picture-thinking". Are teachers and pupils sufficiently aware? Mainx (1955) states:

> Although we are again and again compelled by the high degree of complication of organic events deliberately to simplify by conserving single causal connections outside their relation to the system, and although it is advisable to use language illustratively in the exposition of such connections, still we must always be aware of the dangers which this involves. It is perhaps more difficult in biology, than in the other natural sciences to be popular without being misleading.

Mainx says that many formulations are mere "illustrative circumlocutions" for the assertion of a law, or of a set of conditions, which are testable in experience.

Do teachers and pupils recognize an "illustrative circumlocution" when they are confronted with one? Beckner (1959) says: "My own view is, that philosophers have greatly overestimated the amount of illicit suggestions embodied in teleological formulation. No doubt *some* people are *some*times misled into unverifiable or even false attribution of psychic activity to organic systems". Beckner refers to philosophers, but what about the not-so-philosophically trained high-school pupil?

Textbook writers (university professors writing for upper secondary-school pupils, university undergraduates and *teachers*) while trying hard to be popular, frequently employ "illustrative circumlocutions", a few of which will be quoted below:

> "Why should amphibians have developed these limbs and become potential land-dwellers? . . . . Land limbs were developed in order to reach the water, not to leave it!" (Romer, 1942).

> "Thus terrestrial creatures, although they are not totally able to prevent water-loss, have devised mechanisms that minimise this loss." (Burnett and Eisner, 1964)

> "Natural selection does its best to favour the production of genetic codes guaranteeing behaviour, that increases fitness." (Mayr, 1961).
>
> "For example, if the air temperature outside a human body goes up, the person will make a series of adjustments, including perspiring, to maintain the temperature at a constant level." (Savage, 1964).

The authors of these quotations (representative of a collection of many more) seem to take their readers' sophistication in critically analysing "popularised" statements for granted. Are they justified in doing so?

Research study

The following report embodies some of the data collected within the framework of a project conducted in Israel during 1972–1974. The purpose of the research was to ascertain if, and in what measure, secondary-school pupils (in the 4th and 6th forms) had acquired certain biological concepts, and were able to differentiate between the philosophical assumptions exemplified in a series of biological propositions relating to these concepts. For the purpose of this paper only those propositions relating to the problem of anthropomorphism will be cited. The items to be quoted are from the third and final edition of a battery of four tests I constructed for the study.

Each test was administered to ten 4th form and ten 6th form classes comprising a representative sample of Israeli secondary schools. (In Israel compulsory school terminates with the end of the 3rd form, so the population under discussion must be regarded as an academically gifted group.) The tests were entitled: "Test on the Understanding of the Language of Science" (TULS, Forms A – D).

The following instructions were given to pupils in Test A:

> In this booklet you will find five groups of sentences consisting of four sentences each, all of which describe, or explain the same biological situation or process, but in different ways.
> Read *all four* sentences very carefully, and then decide which of them you might want to use, if you were writing an examination paper on that particular topic.
> On your answer sheet indicate if you would use or would not use *each* sentence, and explain clearly why you would (or would not) use it.
> *Please Note:* You may find all of the four sentences suitable. If so, please indicate which would be the best. On the other hand *none* of the sentences may appeal to you. If so, please state your reasons, and, if possible, propose a sentence of your own.
> *CAUTION* Please do not give reasons relating to style and grammar for your choices. Judge the sentences ONLY according to whether they explain or describe the situation (or processes) scientifically.

Two sample items are given below. It will be seen that each consists of one factual statement, and three others embodying anthropomorphic and/or teleological notions.

> *Item No. 1*
>
> (a) Certain plants, which grow in hot and dry regions, try to develop features which will enable them to reduce water loss.
> (b) Certain plants, which grow in hot and dry regions have features, which enable them to reduce water-loss.
> (c) Certain plants, which grow in hot and dry regions provided themselves with features which enable them to reduce water loss.
> (d) Certain plants grow in hot and dry regions and for that reason features developed in those plants which enable them to reduce water-loss.

Results (% of pupils who *would* choose statement for exam-paper)

	(a)	(b)	(c)	(d)
4th form (N = 20)	15	$\frac{48}{9}$	53	69
6th form (N = 129)	4	$\frac{44}{12}$	47	60
Comparison group of univ. students (3rd year agric. science) (N = 48)	2	$\frac{67}{33}$	38	41

Note: The "fractions" are to be read thus: e.g. 48% chose the statement, but for only 9% it was their only choice.

Item No. 5

(a) When certain primitive plants became too big, they started to develop transport systems in order to distribute water and nutrients more efficiently.
(b) We find, that primitive plants above a certain size, have transport systems, which enable them to distribute water and nutrients more efficiently.
(c) When primitive plants had reached a certain size, additional growth would have raised the problem of efficient distribution of water and nutrients, so they were provided with transport systems.
(d) In certain primitive plants mutations (genetic changes) took place, which were designed to provide them with transport systems for water and nutrients, and thus enable them to attain a greater size, because now distribution was more efficient.

Results	(a)	(b)	(c)	(d)
4th form	35	$\frac{40}{12}$	41	44
6th form	15	$\frac{46}{25}$	35	51
Univ. students	13	$\frac{67}{49}$	23	26

Attention is drawn to choice 1(a) as compared with choice 1(c) and choice 5(a). It would appear that crude anthropomorphism was rarely accepted, whereas "more sophisticated" anthropomorphism turned out to be much more inviting.

Do pupils choose anthropomorphic statements naively, taking the formulation literally, or do they accept it in the metaphorical sense only. TULS – Form D was constructed in order to answer this question. This test was administered to a parallel sample of pupils. The instructions to pupils for Form D were as follows:

> This test contains five groups of sentences. Every sentence in each group describes or explains the same biological situation or process, but each of them does it in a different way. The first sentence is marked "original" and the others: Sentence A, sentence B and sentence C. Read *all four* sentences very carefully, and then decide which of the sentences (A, B, or C) has the same scientific meaning as the original. On your answer-sheet mark the letter representing the chosen sentence and *explain* very shortly the reason for your choice.
> If, *after* reading all four sentences, you have come to the conclusion that none of the sentences (A, B and C) have the same meaning as the "original" mark your

answer-sheet accordingly and explain your reasons.

Item 1 in Form D parallels item 1 in Form A, and so does item 5, the only difference being, that the "factual statement" was here presented as "original", so that the items will not be cited again here.

Results (% of pupils choosing each option)

	Item 1				Item 5			
	A	B	C	NONE	A	B	C	NONE
4th form (N = 308)	41	3	49	8	13	20	30	37
6th form (N = 168)	36	1	48	14	9	15	27	48

It can be seen, that the percentage of pupils attributing the *same scientific meaning* to "sophisticated" anthropomorphic statements (1a, 5c) as to factual statements was quite high, but not so for "crude" anthropomorphisms. Form D thus confirms the results of Form A.

Form B, administered to another parallel sample of pupils, endeavoured to approach the problem from a different angle. The Instructions to pupils were as follows:

> This text consists of 20 sentences, each of which describes or explains a situation or process. Assume that a scientist has been asked to read these sentences, in order to decide if they should be included in a text book on high-school biology. Would he be prepared to accept them in their present form, or would he ask for some changes?

Results: (Mean percentage of pupils *approving* sentences in each category)

	Man	Animals	Plants
4th form (N = 261)	44.4	58.1	50.4
6th form (N = 191)	35.5	49.7	48.4

The high percentage of pupils accepting these formulations as suitable for inclusion in a science-text for schools on the grounds that they "scientifically explain the situation or process" is appaling. The figures quoted above indicate that more pupils are aware of the fact that they have no conscious control over many of their body-functions, than would say the same of plants, and, especially, animals.

The fourth test of the battery, Form C, was developed to test pupils' analytical *and* synthetic facility. It was administered to yet another sample of pupils. Instructions to pupils were as follows:

> In this test you will find five pairs of sentences. All the sentences either explain or describe a situation or process. In *sentence A* of EACH pair the descriptions or explanations are scientifically *un*acceptable, while sentence B of each pair IS acceptable.
> Please read *all five pairs* of sentences, and THEN decide why the sentences marked "A" are scientifically unacceptable. There is ONE COMMON reason for all of them. Find it, and write it on your answer-sheet.

One item-pair will be quoted below:

1.A Deciduous trees, being interested in protecting themselves from winter-frost damage, shed their leaves before winter.
1.B Deciduous trees, which shed their leaves before winter, are thus protected from frost-damage.

In addition to being required to identify the common denominator disqualifying sentences A from being acceptable explanations, pupils were also required to apply this denominator to original formulations of their own, i.e. constructive mode as contrasted with analytical mode. Instructions were as follows:

> You will now be presented with a further two pairs of sentences, however in pair No. 6 sentence A is missing, and in pair No. 7 sentence B is missing. Having carefully read the previous examples (pairs No. 1 – 5), you should have little difficulty in writing the missing sentences yourself. Please do so on your answer-sheet.
>
> 6.A Read sentence B and then write your suggestion on the answer-sheet.
> 6.B When plants suffer from lack of water, the stomata close, and as a result the danger of drying out is reduced.
> 7.A Whenever the reserves of glucose in the muscle of an animals have been used up, it is the liver which despatches replacements there.
> 7.B Compose this sentence yourself and write it on the answer-sheet.

Results

	4th form	6th form
A. % of pupils identifying common denominator AND correctly formulating BOTH sentences	43.3	60.2
B. % of pupils identifying common denominator and correctly formulating ONE of the sentences	17.8	10.0
C. % of pupils identifying common denominator BUT UNABLE to form sentences	3.4	0.6
TOTAL % of pupils identifying common denominator	64.5	70.8
TOTAL % of pupils UNABLE to identify common denominator	35.5	29.2

The percentages of pupils to whom the concept of anthropomorphism is apparently meaningless (pluss category C above), approaches those which in Form B gave their approval to "scientific explanations" containing anthropomorphic formulations.

Discussion

The "problem of anthropomorphisms" was singled out for testing as an exemplar of the much wider class of analogies-as-teaching-instruments. It was shown that the percentage of secondary-school pupils in selective populations who were reay

(a) to make use of anthropomorphic formulations in an examination paper,

(b) to approve of such formulations for inclusion in school biology texts as "scientific explanations";

(c) to equate the meanings of anthropomorphisms with those of factual scientific statements, and

(d) who were unable to identify anthropomorphism as the common denominator in a series of statements *labelled* as non-scientific.

had reached alarming proportions, indeed such proportions which might justify the term "emergency".

Science education journals have rarely dealt with this problem. There were a few papers in the early fifties (Dennis 1953, Crannell 1954, Oates 1956) pertaining to animistic thinking among college students, and reporting alarming results. To my knowledge, there were no reports on these aspects of science education within the last 15 years, till the warning by Hughes (1973) against anthropomorphism. That paper drew an acrimonious rejoinder by Sharefkin and Ruchlis (1974) pointing out that it may be correct to avoid anthropomorphisms in the upper secondary school, but their avoidance in the elementary school must be regarded as a kill-joy, which would prevent the development of further nature-interests of children so deprived. This attitude might be compared with that of the West German Association of Biologists, who, in their proposed biology curriculum of 1973 advocate for *grades 1–4* the gradual reduction of anthropomorphisms, and for grades 5–6 refer pointedly to the dangers inherent in the misinterpretation of phenomena like "insightful" behaviour of animals.

Children enter school "heavily burdened" with anthropomorphic interpretations of their environment. Indeed "in the life of young children in industrial cultures one of the principal functions of anthropomorphized animals is to serve as a medium for the transmission of values, and this function overshadows the real aspects of animals" (Henry, 1971). Be that as it may, there can be no justification for the continuance of anthropomorphic views on the part of biology pupils in the upper secondary school. Only constant attention to the "problem of anthropomorphism" by teachers, teacher-education staff, textbook and "popular" science book writers can prevent results of the type shown in this paper.

It is, however, the teacher, as the only day to day contact person with the pupil, on whose shoulders rests the main responsibility. His own verbal formulations, and his reactions to his pupils' formulations, both oral and written, will determine whether there will be a gradual elimination of pupil-anthropomorphisms.

"Let the teacher beware – cave circumlocutiones illustrativas!"

References

BECKNER, M. *The Biological Way of Thought*. University of California Press, Berkley, 1959.

BRAITHWAITE, R.B. *Scientific Explanation*. University Press, Cambridge, 1953.

BRONOWSKI, J. *Science and Human Values*. Harper & Row, New York, 1959.

BURNETT, A.L. & EISNER, T. *Animal Adaptation*. Holt, Rinehart & Winston, New York, 1964.

CRANNELL, C.W. "The responses of college students to a questionnaire on animistic thinking." *Scientific Monthly*, January, 54-56, 1954.

DENNIS, W. "Animistic thinking amonst college and university students." *Scientific Monthly*, April, 247-249, 1953.

GHISELIN, M.T. "On semantic pitfalls of biological adaptation." *Philosophy of Science*, **34**: 147-153, 1967.

HEMPEL, C.G., & OPPENHEIM, P. "Studies in the logic of explanation." *Philosophy of Science*, **15**: 135-175, 1948.

HENRY, J. "A cross-cultural outline of education." *Cultural Anthropology*, **1**(4), 1960. Reprinted in Henry, J., *Essays on Education*, Penguin, Harmondsworth, 1971.

HUGHES, A. "Anthropomorphism, teleology, animism and personification: why it should be avoided." *Science and Children*, **10**(7): 10-11, 1973.

MAINX, F. *Foundations of Biology*, International Encyclopedia of Unified Science. University Press, Chicago, 1955.

MAYR, E. "Cause and effect in biology", in Lerner, D. (ed.) *Cause and Effect*. Free Press, Glencoe, 1965.

NAGEL, E. *The Structure of Science*, Routledge and Kegan Paul, London, 1961.

OATES, M.E. "Explanation by college students." *Science Education*, **41**: 425-428, 1956.

ROMER, A.S. *Man and the Vertebrates*, University Press, Chicago, 1942.

SAVAGE, J.M. *Evolution*. Holt, Rinehart & Winston, New York, 1963.

SHAREFKIN, B.D., & RUCHLIS, H. "Anthropomorphism in the lower grades." *Science and Children*, **11**(6): 37-40, 1974.

LOGICAL CONNECTIVES IN SCIENCE : A PRELIMINARY REPORT

P.L. Gardner

> Now suppose that we spin the stroboscope so that the slit goes all the way around in exactly the time of rotation of the turntable. Then, each time that we can see through the slit the arrow on the turntable will be in the same position. It will appear to stand still even though it is really rotating. In this instance, then, the time for one rotation of the stroboscope measures the time for one rotation of the turntable. On the other hand, if the stroboscope spins faster than the turntable, the arrow on the turntable will not get all the way around between glimpses, so it will not seem to stand still. Also, if the stroboscope goes too slowly, the arrow will move around by more than one rotation between glimpses; so again it will appear to move. Consequently, by adjusting the speed of the stroboscope to make the arrow appear to stand still, we automatically set the times of rotation equal; and we can use the stroboscope speed – at our control – to measure an unknown time of rotation.
>
> Physical Science Study Committee, *Physics*, (2nd edition), p. 12-13.

The style of writing in this passage is typical of many modern school science text books. A situation is described; a lengthy chain of logical argument is presented; a conclusion is reached.

In this passage, numerous linking words appear which tie the propositions together into a coherent argument. Examples include "Now . . . . so that . . . ." (Sentence 1); "Then . . . . " (Sentence 2); ". . . . even though . . . . " (Sentence 3); "In this instance, then, ... " (Sentence 4); "On the other hand, if . . . . so . . . . " (Sentence 5); "Also, ... " (Sentence 6); "Consequently . . . . " (Sentence 7). Such linking words and phrases may be called *logical connectives*. This paper reports the preliminary stages of an investigation of the difficulties that secondary school students have with this kind of language.

Background to the project

In one sense, the present project simply represents a continuation of my earlier studies in the field of vocabulary difficulties amongst high school science students (Gardner, 1971, 1972a, b, 1974). The central concerns of these studies were to compile a list of non-technical words likely to be encountered frequently in the learning of science, and to investigate students' difficulties with these words.

However, by itself, this interest in the role of language in science education would not have led to the present project. A crucial contribution was made by Strevens (1972) who, in an analysis of technical, technological and scientific English, identified a class of terms which he called *logico-grammatical items*. Strevens divided this class into seven groups, and offered a few examples of each. His list of examples has been reproduced in Gardner (1974).

Strevens does not offer a precise definition of the term 'logico-grammatical item'. However, it is evident that all the terms in his list occur in reasoned discourse. Most (but not all) are terms which act as link words or phrases, tying together clauses, sentences and paragraphs, it is these types of terms which we are here calling *logical connectives*.

A formal definition of logical connectives

Three approaches to defining logical connectives can be adopted.

The first involves making a distinction between logical operators and logical connectives. The term 'logical operator' is used by logicians to refer to words or symbols which link propositions, and at first sight it would seem that logical operators and logical connectives are one and the same. However, there is a difference, in that our class of logical connectives is much broader. The term 'logical operator' usually refers to a very limited number of words such as 'and', 'not', 'or', 'not both', 'either . . . or', 'if . . . then', and 'if and only if . . . then'. Logical operators may be regarded as words which stand for fundamental conceptual rules. However, the category of logical connectives, although it includes all the logical operators, is much broader, for two reasons. First, the English language is so rich that there are many words and phrases which can be used to express those rules. 'Furthermore' is a variant of 'and'; 'alternatively' is equivalent to 'or'; 'suppose that' means 'if' and 'consequently' is equivalent to 'then'. Thus, one reason the class of logical connectives is broader than the class of logical operators is simply that the former class includes various synonyms for logical operators.

But there is another reason, as well. Many terms in English serve to link ideas even if they are not strictly synonyms for logical operators. Terms such as 'actually', 'but', 'for example', 'generally' and 'in short' all serve to connect together propositions in reasoned discourse, yet are not simply reducible to logical operators of the 'and', 'if' and 'not' variety. This project takes a broad view and includes any term which serves to link together propositions within a passage of reasoned discourse, whether the term is an exact synonym for a logical operator or not.

The second approach to the problem of definition draws upon the method of logical analysis, which arose from the 'logical-atomist' school of Bertrand Russell and Ludwig Wittgenstein. Logical analysis involves the reduction of sentences to propositions in their simplest form.* According to the logical atomists, there are two classes of factual propositions: *analysable* and *unanalysable*. Unanalysable ('atomic') propositions are made up of simple words i.e. words which we understand because we are directly acquainted with their referents. For example, the proposition 'This (is) red' is an atomic proposition.

Analysable ('molecular') propositions, on the other hand, are truth-functions of atomic propositions; they consist of atomic propositions linked together by logical connectives. For example, 'This (is) red *and* this (is) square' is a molecular proposition; the word 'and' is a logical connective.

Although this may be helpful, it needs to be stressed that the kinds of propositions encountered in ordinary discourse are hardly ever atomic propositions in the logical atomists' sense of the term. Consider the proposition 'Photosynthesis occurs during the day': this is a *simple* proposition, inasmuch as it asserts one thing only, but it is not *atomic*, because its words are not directly associated with concrete referents. Similarly, the proposition 'Photosynthesis occurs during the day whereas respiration occurs during the day and night' is complex, in that it is made up of two simple propositions joined by a logical connective, but it is far more complicated than the illustrative molecular proposition given earlier. Thus, in order to be useful for our purposes, it is necessary to express our definition of logical connectives in terms of their linkage function between simple propositions, whether or not those propositions are atomic.

The notion that logical connectives link simple propositions is important, in that it allows us to exclude from consideration many terms which resemble logical connectives, but actually occur *within* propositions, rather than as linkages *between* them.

Consider the propositions 'Photosynthesis occurs during the day' and 'A synapse is com-

---

* I am grateful to Ron Laura for helping to clarify some of the ideas in this section.

monly formed by contact of the tip of a terminal branch of the axon belonging to one nerve-cell with the cell-body or with a dendrite of the other nerve-cell'. In these propositions, the words 'during', 'by', 'of', 'belonging to' and 'with' certainly act as linkages, but they link concepts within a proposition: they do not link propositions. Words used in this way (usually, prepositions) have been excluded from consideration in this project.

To summarize the discussion up to this point, then, a term is a logical connective if it links together two complete propositions, but not if it merely links together concepts within a single proposition. This apparently neat dichotomy, however, leaves open the question of what to do with terms which act as a link between a complete proposition and a concept or group of concepts which is not by itself a complete proposition. Two examples will illustrate this intermediate class:

- 'A number of important agricultural pests belong to the order of *Coleoptera*, for example the bollweevil and the wireworm.'
- According to the textbook, the melting point of the salt is 250°C.'

In these sentences, the terms, 'for example' and 'according to' each link one complete proposition with ideas that are not propositions as they stand. For the purposes of the project, we have decided to include such terms in our list of logical connectives. A justification for this is that such terms serve the grammatical function of abbreviating complete propositions. The sentences could be rewritten as

- 'A number of important agricultural pests belong to the order of *Coleoptera*. The bollweevil and the wireworm are agricultural pests which belong to the order of *Coleoptera*.'
- 'The melting point of the salt is 250°C. This assertion is stated in the textbook.'

In both cases, we now have two complete propositions. Clearly, the use of 'for example' and 'according to' results not only in their linkage, but also in their abbreviation at the same time. Consequently, we are now led to define a logical connective as a term which links together a complete proposition with another idea, the idea being either a concept or another complete proposition.

The third approach to a definition of logical connectives is to use grammatical, rather than logical, analysis. We can begin with a broad division of English words into *content* words (those which convey meaning by themselves) and *function* words (those which give grammatical structure to propositions). (Ther terminology is derived from Hartmann and Stork, 1972.) Clearly, logical connectives belong to the latter category. The class of function words includes sub-classes such as articles ('a', 'the') and prepositions ('on', 'to'); the sub-classes to which logical connectives would appear to belong would be *conjunctions* and *conjunctive adverbs* (or *adverbial conjunctions*). Examples of the latter sub-class include 'alternatively', 'consequently', 'moreover' and 'therefore'. Conjunctions are usually classified into two groups: *co-ordinating conjunctions* ('and', 'or', 'but') and *subordinating conjunctions*. This latter group is further subdivided into *qualifying conjunctions* ('if', 'because', 'although'), *prepositional conjunctions* ('as', 'than') and *incorporating conjunctions* ('that', 'what', 'how'). These last two types generally link concepts *within* propositions, but do not form links *between* propositions, and are not therefore included amongst the set of logical connectives. In short, logical connectives would appear to belong to three groups of function words: co-ordinating conjunctions, qualifying conjunctions, and adverbial conjunctions.

By integrating the logical and the grammatical approaches we can form a single definition:

> *Logical connectives are co-ordinating, qualifying or adverbial conjunctions that are used to link a simple proposition with another idea (either a proposition or a concept) to form a complex proposition.*

Research on the learning of logical connectives: a brief review

We can begin to see the relevance of these ideas to education by briefly referring to the works of those cognitive psychologists who are interested in the learning of conceptual rules. In fact, 'conceptual rules' and 'logical operators' are virtually synonymous terms.

Cognitive psychologists have adopted two major approaches in their investigation of how people learn conceptual rules. One is the strictly experimental approach, exemplified by the classic work of Bruner, Goodnow and Austin (1956); the second is the more clinical approach of the Piagetian school.

Studies of concept learning of the Brunerian school have generally employed symbolic stimuli (e.g. nonsense syllables, or geometric figures), and categorization responses. Although these studies have not been directly concerned with the interpretation of meaningful verbal information, they have served to draw attention to the difficulties that learners have with conceptual rules such as negation, inclusive disjunction, exclusive disjunction, conditionality ('if . . . then') and biconditionality (if and only if . . . then'). For a recent study relevant to the present project, see Paris (1973). Studies of the Piagetian school have been much concerned with the sequence of development of logical thinking in children; a child's ability to operate with logical rules such as conditionality is viewed as a specific instance of a more general class of mental abilities which Piaget terms formal operational thinking.

Cognitive psychologists are not, however, the only empirical researchers interested in logical language. Researchers concerned with analyses of the language of teachers and pupils provide another source of relevant research. Writers such as Barnes (1971) have pointed out that different groups of people may employ distinctive linguistic registers, a term which includes vocabulary, sentence construction, degrees of formality of language and modes of reasoning. If a communicator (e.g. a teacher or a textbook author) and his target (pupil) use differing registers, effective communication may well be blocked. This raises the possibility that difficulties with logical connectives, if found, may be an outgrowth of incompatible linguistic registers.

Some recent British research, although not specifically concerned with logical connectives, is worth mentioning here, and that is Bernstein's (1971) well known work on language and social class. Bernstein studied the differences between middle class and working class boys in their use of various grammatical elements, using transcripts of tape recordings made in discussion groups. No class differences were found for the proportion of finite verbs, nouns, prepositions and adverbs, but the middle class *Ss* tended to use a greater proportion of verbs in the passive voice, adjectives and uncommon adverbs.

Bernstein did not obtain data on logical connectives *per se*, but he did obtain data on conjunctions, which form a sub-set of the group of logical connectives. No class differences were found in the *overall* use of conjunctions, but when common conjunctions ('and so', 'or', 'because', 'also', 'then', 'like') were excluded from consideration, it was found that middle class *Ss* did use a higher proportion of uncommon conjunctions. As already stated, middle class *Ss* also used uncommon adverbs more frequently; it will be recalled that conjunctive adverbs form another sub-set of the group of logical connectives. It follows that one would expect to find class differences in the usage of (and hence the difficulty with) logical connectives.

One final investigation worthy of mention is described within a recent review by Brown (1975) of the research on cognitive preferences. Tests of cognitive preference attempt to measure respondents' relative interest in different kinds of information. The pioneer in the field, Heath (1964), developed test items in which an introductory stem is followed by four statements, all of

which are correct. Respondents are asked to select the statement which most appeals to them. The statements reflect preference for *memory* (restatement of facts or terms), for practical *application*, for critical *questioning*, and for fundamental *principles*.

The following item, from Mackay (1971) provides an illustration.

> When a small amount of water is placed in the palm of the hand and allowed to evaporate, the hand is cooled.
>
> A. This effect is used to cool water bags hung on the front of cars.
> B. Heat must be applied to a substance to change its state from liquid to vapour.
> C. No cooling effect would be felt if the water were to evaporate very slowly.
> D. The evaporation of a liquid produces a cooling effect.

In her review of the field, Brown suggests that

> It is possible that pupils, in expressing preferences between statements, are attending to differences in linguistic features (formal characteristics of vocabulary and syntax) rather than differences in meaning.

Brown presents the results of an analysis of 70 cognitive preference items, from three different sources, and she notes that hypothetical or negative words (e.g. 'if', 'would', 'whether', 'not', 'never', 'fail', etc.) occur 56 times in the 70 *questioning* statements, and only three times in the 210 *memory*, *application* and *principles* statements.

Clearly, there is some overlap between logical connectives and Brown's category of hypothetical and negative words. This raises the question of whether those students who are better at comprehending logical connectives are amongst those who have a cognitive preference for statements which include them.

A list of logical connectives

As far as the author is aware, no comprehensive list of logical connectives has ever been constructed. Strevens' (1972) list of logico-grammatical terms contains about a hundred entries (not all of which are logical connectives). Strevens' list does not, however, claim to be comprehensive; in fact, the English language is so rich that several hundred words and phrases can act as logical connectives.

A group of Dip.Ed. students scanned 27 secondary science text and reference books, a total of 2.23 million words, in the four subject areas of physics, chemistry, biology and general science. They recorded frequencies of all logical connectives on a prepared list and added any extra logical connectives that had been omitted. For inclusion in the list each term met one or more of the criteria below.

1. Listed words with a frequency of at last 2 per 100,000 words in the 27 books.
2. Listed words with a frequency of at least 3 per 100,000 words in at least one subject area.
3. Extra words with a frequency of at least 3 per 100,000 words in 2 or more books.
4. Extra words with a frequency of at least 8 per 100,000 words in a single book.

The list of logical connectives produced appears in the appendix.

Measurement of difficulties with logical connectives

Various techniques could be used to measure students' difficulties with logical connectives.

For various reasons, (validity, brevity, and ease of administration) it was decided to employ two kinds of multiple choice items in the project: sentence completion (SC) and gap-filling (GF), illustrated by the following items:

Sentence completion

The weather on the school camp was mostly fine. However

A. some children caught cold after going for a swim.
B. it rained on the day of the hike.
C. we did not enjoy doing all the dishes.
D. there was one really warm day.

Gap-filling

Susan cannot play a musical instrument; _____she likes to listen to the music of others

A. similarly
B. because
C. in effect
D. however
E. in other words

The SC format requires the student, after being given a particular connective, to be able to select a proposition which reflects the logical force of that connective; the GF format, on the other hand, presents the student with a pair of propositions, and requires the student to select the appropriate linking term. The objectives tested by these two formats are complementary, but need not be assumed to be identical.

Pilot project 1. Two pilot projects were conducted late in 1974. The objectives of the first were to establish that logical connectives did, in fact, prove difficult to a substantial number of secondary school students, to provide comparative data on sentence-completion (SC) items and gap-filling (GF) items, and to obtain some cross-sectional data on students at various age levels.

Twenty-five logical connectives listed in Table 1, were selected for testing. The selection of these particular terms was not based on any specific rationale. A 50 item test was constructed, containing 25 SC items followed by 25 GF items.

TABLE 1

Item facilities

Logical connective	Item type	Pilot Project 1					Pilot Project 2	
		Form 1	Form 2	Form 3	Form 4	Total	Mode I	Mode II
however	SC	61	63	75	63	65	53	61
	GF	65	72	75	66	69	87	91
whereas	SC	31	46	42	66	45	43	67
	GF	63	80	81	87	77	60	45
hence	SC	18	11	22	47	23	40	56
	GF	41	46	64	79	56	52	29
thus	SC	53	83	78	79	72	65	75
	GF	45	41	42	61	47	43	34
furthermore	SC	47	54	56	58	53	39	61
	GF	55	61	81	82	68	57	49

Table 1 continued

Logical connective	Item type	Pilot Project 1					Pilot Project 2	
		Form 1	Form 2	Form 3	Form 4	Total	Mode	Mode
despite	SC	63	70	78	76	71	67	68
	GF	63	76	72	87	74	57	45
provided that	SC	69	72	78	82	74	80	85
	GF	88	87	81	100	89	77	80
therefore	SC	47	50	53	76	56	57	64
	GF	57	70	53	61	60	67	64
yet	SC	53	50	53	74	57	62	80
	GF	90	85	89	95	89	66	57
although	SC	78	54	47	79	65	64	75
	GF	78	85	83	84	82	70	76
since	SC	84	93	83	92	88	82	82
	GF	76	78	89	89	82	62	62
so long as	SC	31	48	58	42	44	87	84
	GF	69	85	83	97	82	73	71
in terms of	SC	45	52	56	68	54	87	81
	GF	24	28	22	37	27	77	84
according to	SC	59	61	72	74	65	74	80
	GF	37	33	50	68	46	78	76
nevertheless	SC	29	35	53	53	41	46	57
	GF	41	48	44	61	48	67	59
consistent with	SC	39	35	67	74	51	55	62
	GF	27	24	56	68	42	43	34
in order to	SC	51	52	42	74	54	48	62
	GF	69	70	75	89	75	58	46
indeed	SC	39	50	61	61	51	70	71
	GF	14	20	14	34	20	64	69
with respect to	SC	43	70	58	65	58	63	64
	GF	31	50	56	79	52	71	72
regardless of	SC	59	67	83	79	71	54	64
	GF	65	83	75	92	78	42	38
consequently	SC	47	39	39	63	47	52	63
	GF	35	39	47	84	50	41	40
because	SC	84	83	83	87	84	61	64
	GF	94	93	97	100	96	64	67

Table 1 continued

Logical connective	Item type	Pilot Project 1					Pilot Project 2	
		Form 1	Form 2	Form 3	Form 4	Total	Mode I	Mode II
instead	SC	86	91	86	87	88	43	50
	GF	86	85	92	92	88	26	13
due to	SC	76	78	75	84	78	70	69
	GF	78	87	94	100	89	69	68
on the other hand	SC	57	57	61	66	60	52	56
	GF	63	63	69	71	66	65	57

Key: SC – sentence completion item
GF – gap filling item
Mode I – SC test preceded the GF test
Mode II – GF test preceded the SC test

The test was administered near the end of the school year to all available students (N = 171) in Forms 1, 2, 3 and 4 (grades 7, 8, 9 and 10) at Yallourn High School. (The school serves a large country town whose principal industries are brown coal and electricity production).

Item facilities (i.e. percentage correct) for each item for each grade level are listed in the middle five columns of Table 1, and test statistics are given in Table 2. The test displays a high level of internal consistency, and, as would be expected, there is a steady rise in mean score with grade level.

TABLE 2

Test statistics, pilot project 1

Grade Level	N	Mean	Standard Deviation	Kuder-Richardson 20 Reliability
Form 1	51	28.1	8.1	.86
Form 2	46	30.5	8.8	.89
Form 3	36	32.4	9.1	.90
Form 4	38	37.3	9.0	.92
Total	171	31.7	9.3	.90

Inspection of the item facilities of the two items testing each connective shows some degree of similarity between the values. In other words, if the SC item is difficult, the corresponding GF item also tends to be difficult; if the SC item is easy, the GF item also tends to be easy. The product-moment correlation between the 25 SC item facilities and the corresponding 25 GF item facilities provides a rough index of the degree of similarity; the value obtained (based on the

data from the total sample) was $r = 0.49$.

The fact that the item facilities are moderately correlated is heartening. All that each SC item and its corresponding GF item have in common is the particular connective being tested. The contexts of the items and the tasks required of the respondent are different, yet there is a degree of similarity in item facility. It seems reasonable to assert that this is associated with specific difficulties with the logical connective itself, and if so, the data help to validate the rationale of the project. Had the correlation been near zero, then the item facility would have been determined by factors such as context and item type. The moderate correlation indicates that one can speak meaningfully of a contribution towards the difficulty of the item which is due to the particular connective being tested.

Pilot project 2. Although the first pilot project yielded some useful and interesting data, there were two ways in which the design of the trial was capable of improvement.

(1) For each connective being tested, the contexts of the SC item and the GF item used to test it were different in the first project. One aim of the pilot work was to compare item types, and so in the second pilot project, the same context was employed for both item types, illustrated by the following items:

Test A (Sentence completion)	**Test B (Gap-filling)**
The weather on the school camp was mostly fine. However	The weather on the school camp was was mostly fine. _____ it rained on the day of the hike.
A. some children caught cold after going for a swim.	A. Similarly
B. it rained on the day of the hike.	B. So that
C. we did not enjoy doing all the dishes.	C. Actually
D. there was one really warm day.	D. However
	E. In effect

(2) In Pilot Project 1, each student responded to the set of SC items followed by the set of GF items. In such a design, there is the possibility of practice effect operating i.e. the GF items might display item facilities which have been artificially elevated as a result of previous practice on the SC items. To control for this, a counter-balanced design was adopted. Two separate 25 item tests were prepared: Test A (SC items) and Test B (GF items). Half the students in the sample took Test A followed by Test B (Mode I); the other half took Test B followed by Test A (Mode II).

In addition to these alterations in design, the Pilot Project 1 item data were considered, and a few changes in the wording of stems and distractors were made in order to improve the items. The same 25 connectives were tested.

The tests were administered to all Form 2 students at Clayton Technical School, which is situated in a residential/industrial area in the south-eastern suburbs of Melbourne. The two tests were administered successively within one class period; within each class, about half the students received treatment Mode I, and the rest Mode II. In one class, the supervising teacher apparently failed to follow instructions and the data for that class (N = 18) were removed. This left a sample of 260, (141 under Mode I, 119 under Mode II).

Item facility values of the 25 items are shown in the last two columns of Table 1, alongside the corresponding values obtained in the first pilot project. Mean scores and reliabilities on the two tests under the two modes of presentation are given in Table 3, the analysis of variance of the

scores obtained from the two tests under the two modes of presentation is presented in Table 4. Means and standard deviations of the item facility values are listed in Table 5, and correlations between sets of item facility values are displayed in Table 6.

TABLE 3

Performance on two test formats

Order of presentation	Mean Score			Kuder-Richardson 20 reliability	
	Test A	Test B	Total	Test A	Test B
Mode I (Test A first)	15.4	15.4	30.5	.90	.87
Mode II (Test B first)	17.0	14.3	31.3	.93	.84

TABLE 4

Analysis of variance*

Source of variation	SS	df	MS	F	p
Type of Test	191.9	1	191.9	5.15	< .05
Mode of Presentation	18.3	1	18.3	0.49	N.S.
Interaction	289.6	1	289.6	7.78	< .01
Within cell	19215.2	516	37.2		

* A two-way analysis of variance using an unweighted-means procedure to allow for unequal cell frequencies (Winer, 1962) was employed to analyse the data.

I am grateful to Myint Thien for doing the calculations for this analysis.

TABLE 5

Means and standard deviations of item facilities

Order of presentation	Test A		Test B	
	Mean	S.D.	Mean	S.D.
Mode I (Test A first)	60.6	13.9	61.4	13.8
Mode II (Test B first)	68.0	9.6	57.0	18.9

Considering the mean scores first, (Table 3) it is evident that there is a practice effect which operates to improve scores on the second test when Test B is done first, but not when Test A is done first. This impression is corroborated by the highly significant ($p < 0.01$) interaction effect shown in Table 4.

When Test A is administered first, the mean item facilities and the range of difficulty of Tests A and B, are almost identical (see Table 5). However, when Test B is administered first, the mean item facility of Test B is depressed (and the items display a wider range of difficulty), whereas the mean item facility of Test A is elevated (and the items display a much narrower range of difficulty).

These data imply that when the same connective is tested within the same context under two different test formats, practice at gap-filling (i.e. recognizing which of several connectives best completes a sentence) improves performance on sentence-completion (i.e. recognizing which of several clauses best completes a sentence which includes a connective). However sentence-completion practice does not improve gap-filling ability.

This finding is not too difficult to explain. The student doing Test B first is presented with an incomplete statement (e.g. "The weather on the school camp was mostly fine. _____ it rained on the day of the hike") and must recognize which of several connectives best completes the sentence. When he comes to do Test A, the correct completion is "B. it rained on the day of the hike". His memory of this expression from the earlier test would help him to select this option. For the student doing Test A first, the position is quite different. He is presented with a stem ("The weather on the school camp was mainly fine. However . . . . . .") and must select the best completion. When he comes to do Test B, the correct completion requires recognition of the word 'however'. Commonsense would suggest that the substantive proposition ' rained on the day of the hike' is far more easily remembered than the connective 'however'. This implies that connectives are quickly skipped over by students and are not retained in short-term memory, whereas substantive propositions are more likely to be retained. Again, this piece of evidence provides empirical support for the rationale of the research project, in that it provides a basis for arguing that connectives and substantive phrases are processed by readers in psychologically distinct ways.

Turning now to the correlation data (Table 6), it will be seen that, for both tests, the correlations between the item facilities obtained when it is administered to the random-half samples are high ($r_{A1A2} = 0.87$, $r_{B1B2} = 0.95$). The high values indicate the stability of the item facility indices across different samples of students within the one school. The lower value in the case of Test A is consistent with the earlier discussion on practice effect. When Test B is administered first, this affects students' subsequent performance on Test A; the item facilities of A2 are changed, thereby lowering the correlation between A1 and A2. On the other hand, when Test A is administered first, this does not affect students' subsequent performance on Test B, and the B1/B2 correlation is unaffected.

TABLE 6

Intercorrelations of item facilities

	A1	B2	A2	B1
A1 (Test A administered first)	--	.49	.87	.61
B2 (Test B administered second)		--	.48	.95
A2 (Test A administered second)			--	.54
B1 (Test B administered first)				--

The other four values in the table represent correlations between item facilities of logical

connectives tested by means of two different test formats. As in Pilot Project 1, the values are moderately high (around 0.5 to 0.6), and this gives further support to the proposition that there is some similarity in difficulty between a gap filling item and the corresponding sentence-construction item testing the same word.

Two other sets of statistics are worthy of mention, and these are the frequency distributions of scores and the values of the point biserial correlations, particularly for the SC items. These are shown in Tables 7 and 8. The GF data are included for purposes of comparison.

TABLE 7

Frequency distributions of scores of two tests

Score	Test A (Sentence Completion)		Test B (Gap-filling)	
	Mode I	Mode II	Mode I	Mode II
0-1	0	1	0	0
2-3	1	1	5	2
4-5	4	3	4	3
6-7	13	7	5	6
8-9	13	12	6	14
10-11	22	7	22	14
12-13	13	8	13	15
14-15	11	12	15	14
16-17	12	6	12	17
18-19	8	9	17	11
20-21	10	7	18	14
22-23	8	17	15	7
24-25	26	29	9	2

TABLE 8

Distribution of item discriminations of two tests

Point biserial correlation	Test A (Sentence Completion)		Test B (Gap-filling)	
	Mode I	Mode II	Mode I	Mode II
.10-.19	0	0	0	1
.20-.29	1	0	3	7
.30-.39	6	1	8	9
.40-.49	5	8	11	7
.50-.59	12	11	3	1
.60-.69	1	4	0	0
.70-.79	0	1	0	0

Consider the Test A Mode I data (these are uncontaminated by the practice effect due to

doing Test B first). It will be seen that the frequency distributions of the scores is bimodal: one peak occurs near the perfect score of 25, the other around 10, a few points above chance level. It will also be seen that most of the point biserial correlations for the SC items exceed 0.50. These values are extraordinarily high; indeed, in 15 years experience at constructing and analysing objective tests, I cannot recall ever seeing a classroom test with such a set of characteristics. The data imply that the sentence-completion items measure a highly internally consistent cognitive skill (despite their differences in item content) and that furthermore this is an almost all-or-none skill (i.e. students have either mastered it, or do fairly badly). If one might be permitted some wild speculations at this early stage, one might put forward the hypothesis that the underlying cognitive skill is the ability to process separate simple propositions and re-assemble them into complex propositions. Ability to use logical connectives would seem to be a critical requirement in the utilization of such a skill.

It hardly needs to be stressed that this discussion is built upon data from only one school, and therefore invites caution. Nevertheless, there is clearly a basis for much further investigation here.

Plans for the future

The objectives of the project are four-fold:

(1) to compile a comprehensive list of logical connectives in the English language,
(2) to determine the frequency of occurrence of these terms in scientific material;
(3) to develop a set of objective test items to measure secondary school students' difficulties with the more common terms; and
(4) to investigate possible relationships between the ability to comprehend logical connectives and other variables of interest to education.

Work on the first three objectives is under way. It is intended that each logical connective will be tested in at least four ways: in a scientific context and in an everyday context, and for each context, by means of both a sentence-completion and a gap-filling item. After trial, the final versions will be administered to students at various age levels in a properly constructed sample of Victorian secondary schools during 1976.

The fourth objective is a long term one. One interesting feature of research into difficulties with logical connectives is the potential linkage with other fields of research in science education. Five areas immediately spring to mind: research into cognitive development of the Piagetian variety, research into the retention, processing and transfer of verbal information, exemplified by the recent writings of White & Gagné (1974); research into linguistic registers and the communication gulf between teachers and pupils (see, for example, Barnes, 1971); research into cognitive components of cognitive preferences; and research into relationships between the cognitive and attitudinal aspects of reading scientific material. Questions worth investigating might include the following: Is there a relationship between the ability to comprehend logical connectives and the ability to carry out formal operational thinking? Is the retention of a proposition incorporating a connective affected by the particular connective employed? Are there differences amongst the linguistic registers of textbooks, teachers and pupils with respect to the usage of, logical connectives? Do students skilled at comprehending logical connectives tend to prefer statements in cognitive preference tests which include logical connectives? Does the frequency of usage of logical connectives in a textbook affect pupils' attitudes to that textbook?

Conclusion

To sum up, logical connectives appear to form an identifiable class of terms; it appears to be possible to measure the ability to comprehend them reliably and validly, and it seems possible that this ability might be related to a number of other variables of fundamental interest to science education.

Acknowledgements

I would like to thank my research assistant (Susan Taylor), my typist (Eileen Badrian) and Lourdes Carale, Natalie Drandic, Doris Page, Max Bishop and Janine Ochsenbein (Diploma in Education students at Monash University in 1974) for their assistance in compiling an initial list of logical connectives. Mr Bishop and Mrs Ochsenbein also constructed and administered the tests employed in the pilot testing late in 1974. I am also grateful to the staff and students of Yallourn High School and Clayton Technical School, in particular Mr Frank Dando, for their co-operation during the administration of these tests.

The initial list of logical connectives was converted into the list presented in this paper with the help of a large group of Monash Dip.Ed. students in May 1975. I am grateful to all of them: Julie Barclay, Sue Belcher, Robin Birch, Lyn Breen, Annette Emmerson, Warren Fineberg, Anne Gallace, Frank Gillan, Val Hilton, Jenny Hull, Julie Lee, Paulette Lester, Robyn Mann, Wendy Marshall, June Morris, Heather Nichol, Zoe Nijam, Phil Robertson, Eslyn Shutie, Noel Simm, Lyn Stevenson, Margaret Warren.

This is also an opportune time to thank the South Australian Science Teachers' Association for their invitation to address an in-service course on the topic of language difficulties of science students early in 1974. The act of setting down some of my ideas on paper provided the final, crucial stimulus which converted an idea into an actual research proposal. And finally, of course, I acknowledge the support of the Australian Advisory Committee for Research and Development in Education in funding the project.

References

BARNES, D. "Language in the secondary classroom", in D. Barnes, J. Britton & H. Rosen (eds), *Language the Learner and the School*. Penguin, Harmondsworth, 1971.

BERNSTEIN, B. *Class, Codes and Control*. (Vol. 1: Theoretical studies towards a sociology of language). Routledge & Kegan Paul, London, 1971.

BROWN, S.A. "Cognitive preferences in science: their nature and analysis." *Studies in Science Education*, **2**: 43-65, 1975.

BRUNER, J.S., GOODNOW, J.J. & AUSTIN, G.A. *A Study of Thinking*. Wiley, New York, 1956.

GARDNER, P.L. "Project SWNG: Scientific Words – New Guinea." Report to the Department of Education, T.P.N.G., Melbourne, 1971.

GARDNER, P.L. *Words in Science*, Australian Science Education Project, Melbourne, 1972a.

GARDNER, P.L. "Difficulties with non-technical vocabulary amongst junior secondary school pupils: the Words in Science Project." *Research 1972*, 53-81, 1972b.

GARDNER, P.L. "Language difficulties of science students." *Australian Science Teachers Journal*, **20** (*1*): 63-76, 1974.

HARTMANN, R.R.K. & STORK, F.C. *Dictionary of Language and Linguistics*. Applied Science Publishers, London, 1972.

HEATH, R.W. *Cognitive Preference Test: High School Physics*. Education Testing Service, Princeton, New Jersey, 1964.

MACKAY, L.D. "Cognitive preferences and achievement in physics." Ph.D. thesis, Monash University, 1971.

PARIS, S.G. "Comprehension of language connectives and propositional logical relationships." *Journal of Experimental Child Psychology*, **16**: 278-291, 1973.

STREVENS, P. "Technical, technological and scientific English." Paper presented to the Conference of the International Association of Teachers of English as a Foreign Language, London, 1972.

WHITE, R.T. & GAGNÉ, R.M. "Effect of degree of organization of information on retention." Paper presented at meeting of the Australian Association for Research in Education, Melbourne, 1974.

WINER, B.J. *Statistical Principles in Experimental Design*. McGraw Hill, New York, 1962.

APPENDIX

List of logical connectives

above all . . . according to (= corresponding to) . . . according to (= on the authority of)
actually . . . after . . . again . . . also . . . alternatively . . . although . . . and so . . . apart from . . .
apparently . . . as (= because) . . . as (= when/while) . . . as (= similar to) . . . as a result . . .
as follows . . . as if . . . as long as . . . as much as . . . as shown (by) . . . as soon as . . .
as though . . . as to . . . as we have seen . . . as well as . . . as you know . . . at first . . .
at least . . . at that/the same time . . . at this/that stage . . . basically . . . because . . .
because of the fact that . . . before (long) . . . both . . . but . . . but also . . . but if . . . by . . .
by far . . . by means of . . . by way of . . . certainly . . . clearly . . . commonly . . .
compared (able) with . . . consequently . . . considering that . . . conversely . . .
depending (ent) (up) on . . . despite . . . due to . . . e.g. . . . either-or . . . especially . . .
essentially . . . even . . . even if . . . even though . . . even when . . . eventually . . . evidently . . .
exactly as . . . except . . . finally . . . first(ly) . . . for (= because) . . . for example . . .
for instance . . . for this/that reason . . . formally . . . fortunately . . . frequently . . .
from . . . from the point of view . . . further (= furthermore) . . . furthermore . . . generally . . .
hence . . . here . . . however . . . i.e. . . . if . . . if-then . . . if so . . . in addition (to) . . .
in * case(s) . . . in contrast . . . in * example(s) . . . in fact . . . in general . . .
in * instance(s) . . . in order to . . . in other words . . . in particular . . . in practice . . .
in terms of . . . in that . . . in turn . . . in * way(s) . . . including . . . indeed . . . instead of . . .
it can/was/will/ be seen/shown that . . . it follows that . . . it seems clear that . . .
it would appear/seem that . . . just as . . . just what . . . known as . . . like/much like/
very much like . . . moreover . . . namely . . . neither . . . neither-nor . . . nevertheless . . .
next . . . nor . . . not only – but . . . not only – but also . . . notice (that) . . . now (= instead)
now (= simultaneously) . . . now (= consequently) . . . nowadays . . . obviously . . . occasionally . . .
of course . . . often . . . on the basis of . . . on the other hand . . . once (= if) . . . only . . .
only if . . . or . . . other than . . . otherwise . . . particularly . . . perhaps . . . plus . . . possibly . . .
probably . . . provided(ing) that . . . rather than . . . really . . . recently . . . respectively . . .
say . . . second(ly) . . . see . . . similarly/similar to . . . simply . . . simultaneously . . . since . . .
so . . . so also . . . so as . . . so far . . . so long as . . . so that . . . so to speak . . . somehow . . .
sometimes . . . still . . . such . . . such as . . . such is . . . such that . . . suppose (that) . . .
that is (= in other words) . . . the fact that . . . then (= it follows that) . . . thereby . . .
therefore . . . third(ly) . . . this means/suggests that . . . though . . . thus . . . to (= in order to)
together with . . . too . . . unfortunately . . . unless . . . unlike . . . until . . . usually . . . viz . . .
when (= as soon as) . . . whenever . . . where . . . whereas . . . whereby . . . whether or not . . .
which . . . while (= when) . . . while (= provided that) . . . while (= on the other hand) . . .
whilst . . . yet

* Many variants of these connectives are found, in which the gap is filled by a quantifying adjective (many, one, one such, every, all, some, some such) or a qualifying adjective (the, these, other, general, which, various, similar, following, either, such a, same).

GRADE PLACEMENT OF PHYSICAL AND CHEMICAL CHANGE CONCEPTS IN PRIMARY SCHOOL SCIENCE

John W. Fisher & Robert L. Kimber

Introduction

A number of new sets of curriculum materials in primary school science have been recently forthcoming. Most of these materials have emanated from U.S.A. and U.K. We wanted some indication of the suitability of materials produced overseas to the Australian situation. One method, which is currently being employed in South Australia (by Collings and Fisher), and Tasmania (Rodwell, 1975) is to use commercially prepared resources and evaluate their effectiveness in schools. The other is to determine the optimum grade level at which particular concepts can be attained by students. After so doing, the prepared curriculum materials may be viewed as suitable or not on the basis of the findings: or new materials may be written.

In this study, we attempted to determine how effective a selected method of instruction was in teaching the classificational concepts of physical and chemical change at a number of grade levels. This work fits into the second category mentioned above.

Selection of concepts

As in the work of Pella & Voelker (1968), the seventh conceptual scheme of the NSTA (Voelker, 1964) was selected as the central focus for this research:

> All matter exists in time and space and, since interaction occurs among its units, matter is subjected in some degree to changes with time. Such changes may occur at various rates and in various patterns.

In scientific investigations, the large amount of information that is steadily amassed needs classification in order to become meaningful and useful. Within the larger conceptual scheme of "change", we selected the two lesser concepts of "physical change" and "chemical change" around which a sequence of instruction was established. Two criteria, those of reversibility of an event and formation of a "new product" were used in defining physical and chemical changes. For the purposes of this study, a physical change (P.C.) was defined as one which was reversible and/or in which no "new product" was formed as a result of the change. A chemical change (C.C.) was taken to be one which appeared irreversible and/or in which a "new product" was formed.

Method of instruction

One of the biggest problems with primary school science is in obtaining teachers who have the confidence to teach it. As a consequence of this, most science taught in South Australian primary schools is done in fairly teacher-directed learning situations. Since we did not want the method of instruction to differ greatly from that which the children were used to, we employed a semi-formal approach. The students were given instruction cards which outlined some activities and discussion points for small group work and as a whole class. (See Appendix A.1).

Ridley Grove study

We are grateful to the Principal and staff of Ridley Grove Primary School who made avail-

able a grade 4 class (24 students) and a grade 6 class (28 students) in an open-space unit for the purposes of our study.

Pre-test. All of the children in each of these grades were given a pre-test consisting of five questions:

1. Do you know what the word change means?
2. Have you ever heard of the words PHYSICAL CHANGE and CHEMICAL CHANGE?
3. What do you think happens in a PHYSICAL CHANGE?
4. What do you think happens in a CHEMICAL CHANGE?
5. If you saw a pair of changes and you were told one was a physical change and one was a chemical change, how would you tell which one was the chemical change?

The questions were read to the students and they ticked the Yes, No or Don't Know boxes or completed the sentences where appropriate.

Teaching sequence. Five lessons (1 hour per week) were given to each class. Five sets of experiments, illustrating physical and chemical changes, were done by the students. Each set of experiments contained a pair of phenomena chosen on the basis of three criteria:

1. Each pair was to include one physical and one chemical change phenomenon.
2. The materials used in each pair of experiments, that were illustrative of P.C. and C.C., were to be as alike as possible.
3. The number of clues (physical phenomena such as gas evolution, colour change commonly used as indicators of chemical changes) used by children in making their decision whether the change was physical or chemical was to be a minimum one wherever possible. The clues were present only with C.C. phenomena. The clue (distractor) was visually negated by a corresponding P.C. phenomenon. (See Appendix A.2 for Sets 1–5).

Lesson 1. There was whole class discussion of the word "change" and the role of the modifier word was discussed using examples such as hot change, weather change etc. Demonstrations of change in material objects led to discussion of how changes in matter could be detected by comparing initial and final properties of materials undergoing change.

Lessons 2–4. Children did experimental sets 1–5. In the discussion at the end of each lesson, the teacher emphasised the role of reversibility and new product formation in classifying changes as P.C. or C.C.

Lesson 5. There was a review of change (with examples), differences between physical and chemical change and the role of clues as indicators of P.C. and C.C.

Post-test. There were three parts to the post-test given at the end of the lesson series.

Part 1 consisted of three questions which were printed on a sheet and also read to the children:

1. What do you think a physical change is?
2. What do you think a chemical change is?
3. How would you tell if a change was a physical or a chemical change?

The children wrote their answers to these questions in the spaces provided on the paper.

Part 2 consisted of ten *demonstrations* illustrating physical and chemical changes. The demonstrations were similar to, but not the same as the experiments used in the instruction sets

1–5. After each demonstration the children were required to indicate on an answer sheet whether it was an example of a physical or chemical change, and *why* they had selected that answer.

Part 3 consisted of ten examples, illustrating P.C. and C.C., which were *described* to each child who then decided whether it was a P.C. or C.C. and why.

Results and discussion. Data consisting of scores on the tests were used in giving a measure of the children's level of understanding of the concepts of P.C. and C.C. relative to three arbitrary levels. Indication of the level of concept attainment was given by each student's ability to:

I. define the concepts of physical and chemical change
II. classify observed phenomena as physical or chemical changes
III. support correct classifications with valid reasons.

Pre-test results. Analysis of the responses to the pre-test questions indicated that the students had negligible understanding of change and did not possess any understanding of physical and chemical change.

Post-test results. The proportion of students who correctly answered questions 1–3 are shown below:

	Question 1	2	3
Grade 4	63%	63%	25%
Grade 6	89%	85%	36%

The results indicate that the majority of the grade 4 and 6 children (more Grade 6 than Grade 4) could define a physical change and a chemical change. They could not, however, apply the definition to distinguish between P.C. and C.C.

Table 1 shows that, although both the grade 4 and 6 children could correctly classify more than 50% of the demonstrated and described phenomena as illustrating P.C. or C.C., the grade 6 children scored significantly higher than those in grade 4.

TABLE 1

Ability to classify phenomena

Mode	Demonstrated		Described	
Grade	4	6	4	6
$\bar{x}$*	5.3	7.35	5.87	7.28
S.D.	2.18	1.79	2.36	1.88
U	5.6 $z = 3.3$		454 $z = 2.17$	
SIG.[B]	$p < .001$, G6 > G4		$p < .05$, G6 > G4	

* Maximum mean score = 10.

B $\alpha = .05$, one tailed test. Mann-Whitney U Test.

There was quite good correlation ($p < .01$) between correct classification of P.C. and C.C., for Grade 6, $r_s = .75$, for Grade 4, $r_s = .86$.

TABLE 2

Classifications supported by correct reason

Grade	4	6
$\bar{x}^*$	1.5	5.25
S.D.	2.63	3.02
U	552.5 $z = 3.97$	
SIG.[B]	$p < .001$, G6 > G4	

* Maximum mean score = 10.

B. $\alpha = .05$, one tailed test.

Table 2 indicates that the Grade 6 children have a fair grasp of the concepts of physical and chemical change in that they could support just over 50% of their correct classifications with correct reasons (Level III). In other words, they could apply the definitions or classificatory labels of P.C. and C.C. in just over half the examples demonstrated to them.

Magill study

We are appreciative of the opportunity given us by the principal and staff of Magill Primary School to undertake this study with some of their students. Two grade 4–5[1] and two grade 5–6 composite classes were arbitrarily assigned as experimental (E) and "control" (C) groups, as below:

	Grade 4	Grade 5[1]	Grade 5	Grade 6
Control	18	9	12	11
Experimental	16	5	10	10

The low number of students in grade 5[1] caused some problems in the interpretation of results.

Since streaming does not occur in this school and I.Q.'s were not measured for the students, we based our notion of approximate matching of C to E on the pre-test results.

Pre-test. One drawback we noticed with the Ridley Grove study was that the pre-test did not determine whether or not the students had the ability to recognize the reversible nature of an event or the formation of a "new product" when some materials underwent change.

In this case, the students were given twelve demonstrations illustrating P.C. and C.C. The students had diagrams of the input (A), process (B) and output (C) on the sheet in front of them (see Appendix A.3) as well as being able to watch the materials change. The students were asked if the material in (C) could go back to what it was like in (A). They had to tick yes or no.

We did not want to give the students the "choice" of P.C. or C.C. as being the answer

for the name of the change occurring in the materials. We simply left a space in which they could indicate a special name for the change if they so desired and the reason why the chose it.

Teaching sequence. The same semi-formal method of instruction was employed with the experimental classes in this study as was used with the Ridley Grove children. There were, however, six lessons given in this sequence as it was found that the five given previously were not quite sufficient for the students' apparent needs, i.e. the same experimental sets 1–5 and sequence was employed but one week more was taken for instruction at Magill.

Post-test. A similar format to that in the corresponding pre-test was used. Twelve demonstrations were given and students wrote their responses on the paper after each demonstration. As well as asking about the apparent reversible nature of the events, the children were asked to indicate whether or not a new product had been formed, what special name they would give for the change and why they would use that name.

Retention test. A test identical in format to the post-test was given to the students five weeks after the post-test. This five week break included a fortnight of holidays for the students. The retention test gave us an opportunity to investigate the suggestion by Myer (1962) that an incubation period following intensive study may facilitate learning.

Results and discussion. Indication of the level of concept attainment was given by each student's ability to:

I. (a) identify the reversible nature of an event
   (b) identify the formation of a new product
II. classify observed phenomena as physical or chemical changes
III. support correct classifications with valid reasons.

Pre-test. The results in Table 3 indicate that the students in both the control and experimental grades were quite able to identify the reversible nature of the events demonstrated to them. They did not, however, use the terms P.C. or C.C.

TABLE 3

Pre-test results: Magill

Level		Grades							
		C4	C5[1]	C5	C6	E4	E5[1]	E5	E6
I(a)	$\bar{x}$*	9.3	10.4	9.2	9.8	9.8	10.6	9.3	9.4
	S.D.	1.4	.7	1.1	1.6	1.2	.6	1.3	1.1
II	$\bar{x}$	0	0	0	0	0	0	0	0
III	$\bar{x}$	0	0	0	0	0	0	0	0

* Maximum possible mean = 12

Mann-Whitney U tests between and within the control and experimental classes revealed no significant differences in students' understanding at Level I(a).

Post-test. The results in Table 4 indicate that the students in both the control and experimental grades possess a reasonable grasp of the notion of P.C. and C.C. at the I(a) and I(b) levels. The students in classes E5[1], E5 and E6 scored significantly higher ($p < .05$ Mann Whitney U-test) than the corresponding control classes for Levels II and III.

TABEL 4

Post-test results: Magill

Level		Grades							
		C4	C5[1]	C5	C6	E4	E5[1]	E5	E6
I(a)	$\bar{x}$	7.6	8.1	8.5	7.7	8.3	9.8	8.9	7.9
	S.D.	.9	.6	.5	.9	1.4	.5	1.5	1.1
I(b)	$\bar{x}$	6	6.2	6.5	7.1	6.9	9.4	7.5	7.1
	S.D.	1.2	1.1	1.5	1.5	1.5	.9	1.9	1.2
II	$\bar{x}$	0	.1	.3	0	2.6	9.4	6.2	4.6
	S.D.	0	.3	.65	0	3.6	.9	2.3	2.4
III	$\bar{x}$	0	0	0	0	2.5	9.4	5.7	3.8
	S.D.	0	0	0	0	3.4	.9	2.8	3.1

Maximum possible mean = 12

The Wilcoxon sign test was employed to detect any significant differences which existed in a group's concept attainment level score on pre- and post-tests, where similar items were used. The only significant difference observed in the control groups was for C5; post- being higher than pre-test score for Level I(a) ($p < .05$).

In the experimental groups, no significant differences were obtained at Level I(a). The experimental groups however, displayed post-test attainment significantly higher than pre-test for Levels II and III ($p < .05$). This greater development of understanding of concepts was attributed to the instruction experienced by the children in the experimental groups.

Retention test. Comparisons of the control groups' concept attainment level scores indicate that in all but C5 the groups scored significantly higher ($p < .05$) on the post-test than on the retention test, where similar items were used, for Levels I(a) and I(b).

The only significant difference between post and retention scores for any level within the experimental groups was experienced by E5 with retention-test scores for Levels II and III significantly higher ($p < .05$) than post-test attainment. The experimental groups for grades 5 and 6 appear to have obtained around a 50% level of success in their ability to classify events as P.C. or C.C. and give a valid reason for so doing.

The significant decline in the control groups' attainment which was not evidenced in the pre- to post-test comparison, may indicate possible confusion with the test items arising in these students due to a lack of feedback about their performance on prior tests. The experimental groups did not receive any feedback on their test performances but they did benefit from the instructional sequence.

TABEL 5

Retention test results: Magill

Level		Grades							
		C4	C5[1]	C5	C6	E4	E5[1]	E5	E6
I(a)	$\bar{x}$	5.9	6.1	6.5	6.3	8.1	8	8.8	7.5
	S.D.	1.1	1.4	1.3	1.2	1.2	.7	1.3	1.1
I(b)	$\bar{x}$	3.8	4	4.3	4.9	7.1	8	7.6	6.3
	S.D.	1.2	1.2	2.2	1.4	1.9	.7	2.5	1.5
II	$\bar{x}$	.1	.3	.5	.2	3	4.8	7.4	4.8
	S.D.	.3	.7	.7	.6	3.9	4.4	2.4	2.7
III	$\bar{x}$	0	0	0	0	2.7	3.2	7	4.7
	S.D.	0	0	0	0	3.9	4.4	2.5	2.9

Maximum mean score = 12

Conclusions and implications

Both the control and experimental groups of students in grades 4 to 6 at Magill revealed that they could indicate reasonably well whether or not an event was reversible, (Level I(a)). This pre-test score however, was not a predictor for success at higher levels of understanding of P.C. and C.C. concepts (Levels II and III) for all the experimental groups after the instruction had occurred. The modified pre-test could be used to determine whether or not children have the ability to identify the reversible nature of an event and/or formation of a new product.

This semi-formal method of instruction appears to aid the development of higher levels of understanding of physical and chemical change concepts. It helped some students, notably those in grades 5 and 6, to group or classify their prior knowledge and support the classification of new examples of P.C. and C.C. with correct reasons 50% or more of the time.

The increasing maturity of the students, as indicated by higher grade level and greater experience with examples of P.C. and C.C. phenomena appear to be two factors influencing the attainment of higher levels of the classificational concepts of physical and chemical change.

The students were more able to identify the properties of reversibility and new product formation when simple examples of change phenomena were used in the tests. In items such as breaking a balloon and sharpening a pencil the notion of reversibility, as it applied to classifying the change as a P.C., led to confusion. Accurate discernment of most of the chemical precipate events was difficult for the majority of the students. The difficulty, in terms of concrete and abstract nature, of the items used in the tests is thus very important in determining the student's level of understanding of P.C. and C.C. concepts. (Details of items used in this study may be obtained from the authors.)

A specialist science teacher was used in these studies. Would this method of instruction have been suitable for the normal class teachers? Would they have attained comparable results? Would the same teacher obtain similar results with students in other schools? These are a few of

the many factors which plague researchers who investigate grade placement of concepts such as physical and chemical change in primary school science.

References

MYER, M.E. "Some implications of empirical studies in higher thought processes for science education." University Microfilms, Ann Arbor, Michigan, 1962.

PELLA, M.O. & VOELKER, A.M. "Teaching the concepts of physical and chemical change to elementary children." *Journal of Research on Science Teaching*, **5**: 311- , 1968.

RODWELL, G. "Early childhood science: Tasmania." *Australian Science Teachers Journal*, **21**(*1*): 104-106, 1975.

VOELKER, A.M. *Theory into Action.* National Science Teachers Association, Washington, 1964.

APPENDIX

A.1. Instruction and study guides used for student experiments during instruction Set 1 Period.

Experiment 1.

(a) All take a piece of black paper. Write down its properties and its state.

(b) Tear the black paper into 2 strips.

(c) Are any of the properties or state of the paper different? If they are, write them down.

(d) Has a change happened? Why do you think so?

Talk about the next two questions in your group, before you write an answer down.

(e) Has a new thing (product) been made by this change?

(f) Could the torn paper go back to what it was before the change?

Experiment 2. BE VERY CAREFUL

(a) The spokesman of the group is to pick up a new piece of black paper. All of the group is to write down the paper's properties and its state.

(b) Tear the piece of paper into 2 strips.

(c) The spokesman and another person will be needed to do this part. The rest of the group is to *stand back* and *watch*!
The leader is to clip one strip of paper into the peg, and hold it AWAY from himself and others, keeping it *over and close* to the tray.
The other person may now light the paper. When it has burnt out, allow the remains to fall onto the tray. Repeat this again for the second strip of paper.

REMEMBER TO BE VERY CAREFUL

(d) Are any of the properties or state of the paper different? If they are write them down.

(e) Has a change happened? Why do you think so?

(f) Has a new thing (product) been made by this change?

(g) Could the burnt paper go back to what it was before the change?

A.2. Experiments done by children were as follows –

Set No. 1

Physical Change: Black paper was torn into strips.

Chemical Change: Black paper was torn into strips and burned.

Distractor: Light emission.

Distractor Negation: Set No. 2.

Set No. 2

P.C.: A small light bulb was lit using batteries.

C.C.: A small flash bulb was ignited.

Distractor: Change in form.

Distractor Negation: Set No. 1.

Set No. 3

P.C.: Copper sulphate solution was mixed with water.

C.C.: Copper sulphate solution was mixed with a solution of barium chloride.

Distractor: Precipitate formation.

Distractor Negation: Melted wax was mixed with cold water.

Set No. 4

P.C.: Acidic phenolphthalein solution was mixed with water.

C.C.: Acidic phenolphthalein solution was mixed with basic solution.

Distractor: Colour change.

Distractor Negation: Blue and yellow water paints were mixed.

Set No. 5

P.C.: Common salt was dissolved in water.

C.C.: Sal-vital was added to water.

Distractor: Gas evolution.

Distractor Negation: Soda water was poured into a beaker.

A.3. Pre-test: Answer Sheet used to assess understanding of physical and chemical change concepts.

1.

Questions.

(1) Could the material in C go back to what it was like in A?

ANS. ☐ Yes ☐ No

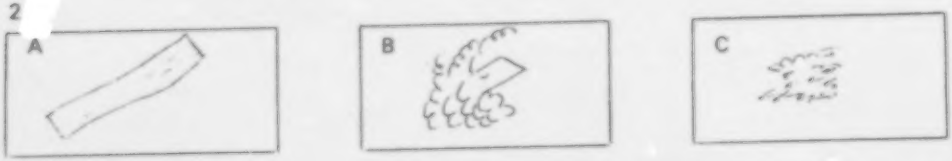
(2) What special name could you use for this change?

ANS. I could call it a ..... change.

(3) Why would you use that name?

ANS. I would use that name because .....

2.



Questions

(1) Could the material in C go back to what it was like in A?

ANS. ☐ Yes ☐ No

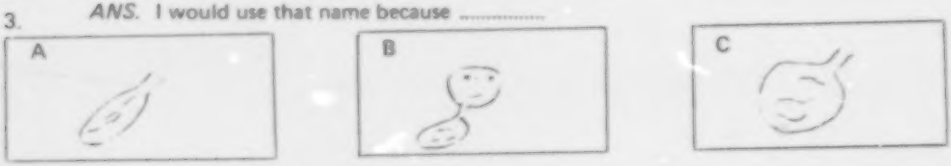
(2) What special name could you use for this change?

ANS. I could call it a ..... change.

(3) Why would you use that name?

ANS. I would use that name because .....

3.



Questions

(1) Could the material in C go back to what it was like in A?

ANS. ☐ Yes ☐ No

(2) What special name could you use for this change?

ANS. I could call it a ..... change.

(3) Why would you use that name?

ANS. I would use that name because .....

HAS PAPUA NEW GUINEA'S THREE PHASE PRIMARY SCIENCE PROGRAM HAD ANY IMPACT AT SECONDARY LEVEL?

M.N. Maddock

The three phase primary science program

Over the last twenty years or so there has been much concern with promoting the teaching of science in schools, with particular interest in its place in the educational programs of developing countries, and UNESCO and UNICEF have been to the fore in promoting moves for the improvement in science teaching. The regional workshop on planning for science teaching improvement in Asian schools, conducted jointly by these two bodies in Bangkok in 1969, stated a policy on the aims of science teaching as follows:

> The aims of education today which can be achieved through the teaching of science can be divided into two groups. They refer to: *the child as an individual*, and the *child relative to his future in society*.
>
> *To help the child attain:*
>
> (a) adaptability to changes in his society;
> (b) self-confidence which involves the ability to make decisions;
> (c) the ability to continue to gain knowledge from surroundings throughout his life;
> (d) a critical and inquiring mind;
> (e) an attitude of co-operation to work with his fellow-men.
>
> *To lead to the following outcomes:*
>
> (a) the development of appropriate scientific knowledge and concepts;
> (b) the acquisition of psychomotor and mental skills;
> (c) development of scientific attitudes.
>
> (UNESCO, 1969, p. 7)

In keeping with this philosophy, the two bodies have supported science teaching programs throughout the Asian-Pacific region. One such supported project is the Three Phase Primary Science (TPPS) Program in Papua New Guinea, now well into its implementation stage.

A committee of teachers and senior officers from the Papua New Guinea Education Department and a science expert provided by UNESCO began working on a first draft in 1969 and by 1971 about 600 schools were teaching the full course. This number has been consistently expanded since. The course was produced in three phases. The Phase 1 (grades 1 and 2) series of simple activities are aimed at providing experiences leading to basic concepts not normally obtained in the Papua New Guinea home situation. Phase 2 (grades 3 and 4) develops simple observational experiments to a further degree and Phase 3 (grades 5 and 6) consists of somewhat more formal experiments involving skills of observation, measurement and recording which are aimed at the formulation of basic scientific principles, and much more structured than the other two phases. Kits of simple equipment in specially designed cabinets were provided during the developmental stage by UNICEF, and supply is being continued by the Papua New Guinea Education Department as the expansion stage of implementation is carried out.

The general aim of the course is:

> To allow the children to gain a knowledge and understanding of the world around them in as interesting a way as possible through activity and enquiry.
> To develop this attitude of inquiry. (Williams, 1971, p. 9)

The activities are specified in detail on special activity cards, each card representing a single lesson.

Evaluation at primary level

In order to obtain an evaluation of the success of the project, the Education Department of Papua New Guinea provided a research grant to Michael Wilson of the University of Papua New Guinea, who has issued an evaluation report on Phases 1 and 2 (Wilson, 1972(a)) and a preliminary report (Wilson, 1972(b)) and interim report (Wilson, 1974) on Phase 3.

He concluded that in many respects, the Phase 1 and 2 project had fulfilled the organisers expectations (Wilson, 1972(a), p. 9). Interest had been generated in the children, but a major failure had been the lack of success in engendering a spirit of enquiry.

However, the preliminary and interim evaluations suggest that Phase 3 has not been so successful. In almost all Phase 3 lessons observed some major difficulty was encountered by teacher and pupils – the most frequent arising from the lack of understanding on the part of the teacher – either on principles underlying the activities or the practical details of the activities themselves (Wilson, 1972(b)).

Such difficulties appear to have resulted in a reluctance to teach Three Phase Primary Science (Wilson, 1974, p. 19).

In the cognitive tests prepared items asking for simple, direct knowledge of a phenomenon included in TPPS were generally well answered by pupils, but those going beyond this to application or development of the basic idea were much less well answered (Wilson, 1974, p. 21). Results on an attitude test indicate that TPPS tends to lead pupils to an exaggerated view of the infallibility of science, although any differences between TPPS and non-TPPS pupils are small (Wilson, 1974, p. 25).

Evaluation of influence on secondary science

Although TPPS had certain specific aims for pupils terminating their formal schooling at the end of grade 6, the developers of the scheme also had a more far reaching aim of improving performance at secondary level.

During 1971 the Secondary Science Committee produced draft materials for the first stage of a new Forms 1 and 2 science course, which was put through preliminary trialling in 1972 and revised and retrialled in 1973. In conjunction with the 1973 trialling, achievement tests were produced and used in the evaluation.

This provided a unique opportunity for further evaluation of the TPPS course, as students in the trial schools included those who had been subjected to TPPS Phase 3 lessons in 1970 and 1971, as well as those who had not done the primary course, a circumstance which would not arise again as TPPS is expanded into all primary schools.

The pilot secondary materails were being trialled in six schools – three in Port Moresby, two in the Eastern Highlands and one in Madang – not by any means a random sample of Papua New Guinean high schools, but reasonably representative. Two of the schools were mission schools and four government.

The pilot teachers were asked to identify students who had been through TPPS at primary schools. As suitable pupil and school records were not available for this task, the teachers were asked to show their pupils the TPPS cards, discuss the matter with them and then categorize them into three groups – those definitely identified as former TPPS pupils, those definitely not, and

those uncertain, and to designate them with a code number in the identification number for the achievement tests.

About 700 pupils with TPPS experience were identified in the pilot schools and 270 in the control schools, with about 60 and 100 non-TPPS pupils respectively. The numbers varied slightly from test to test due to absences from school. Obviously this method was not as satisfactory as classifying from definitive records, but the science committee was satisfied that any error in the categorization achieved was relatively minor; in any case it was the only method available, and the opportunity to obtain some insight into the impact of TPPS was too important to miss. Tests were run on Units 1, 2 and 3 of the revised Pilot Secondary Science Syllabus (Department of Education, PNG, 1972) and the results of these tests were analyzed on the Public Service Board's ICL 1903A computer in Port Moresby using the PNG EDP Item Analysis Package written by one of the Centre's programmers on specifications provided by the author.

This program provided individual results, score frequency distribution, mean, standard deviation and KR (20) reliability coefficient on multiple choice objective tests, together with success indices, point biserial correlations and item-choice frequency distributions for the total population and for designated sub-populations.

In the evaluation of the secondary science materials, three control schools not doing the new pilot course had been designated in and near Port Moresby and the Unit 1 and Unit 3 tests were run on both pilot and control schools, the Unit 2 test being run only in the pilot schools. Ex-TPPS and non-TPPS students were identified in both pilot and control groups.

Comparisons were made between the results of the ex-TPPS and non-TPPS students on each of the Unit 1, Unit 2 and Unit 3 tests. In addition the results on individual items on these tests specifically related to Phase 3 TPPS activities were compared.

Whole-test results. The Unit 1 test was on simple measurement principles which were given considerable attention in the introductory unit of the pilot secondary syllabus, but not in the early stages of the interim syllabus being used by control schools. A number of the items in the test are closely related to the activities specified on lesson cards in both Phase 3A and Phase 3B of TPPS.

The reliability for the Unit 1 test was not high, and some caution needs to be exercised on any comparison between means. In the pilot schools, no significant difference was found between the means of the ex-TPPS and non-TPPS groups, using a t-test for the difference between means (see Table 1). In the control schools the difference was much larger, the TPPS group scoring the higher, but still not significant at the 0.05 level although significant at the 0.06 level (see Table 1). A point biserial correlation between the scores and group membership obtained from the t-ratio (Cohen, 1965, p. 104) was only 0.10, indicating that only a very small percentage of the variance was associated with membership of the ex-TPPS and non-TPPS groups.

The Unit 2 test introduced simple work and energy concepts related to the particle model, and three of the items were closely related to lessons on Phase 3B. The reliability for this test was again not high, and similar caution needs to be exercised on interpreting the difference between means. However, the difference of 0.13 between the ex-TPPS group and the non-TPPS was not significant at the 0.05 level (see Table 1).

The Unit 3 test on the thermal expansion unit did not contain any items directly related to activities in either Phase 3A or Phase 3B. Again no significant difference was detected between the means within the pilot schools, but in the control schools the difference between means of 1.3 was significant at the 0.01 level, with a point-biserial correlation between the scores and group

TABLE 1

Results of whole tests

Test	Sample	Group	N	Mean	SD	KR20	t	p	r_{pb}
Form 1 Unit 1 (Measurement)	Total Population	-	1215	8.55	2.85	0.45	-	-	-
	Pilot Schools	TPPS	710	8.91	2.73	-	0.21	>0.05	-
		Non-TPPS	99	8.71	2.76	-			
	Control Schools	TPPS	274	7.91	2.88	-	1.89	<0.06	0.10
		Non-TPPS	59	7.12	2.88				
Form 1 Unit 3 (Energy & Particles)	Total Population	-	752	10.20	2.84	0.44	-	-	-
	Pilot Schools	TPPS	677	10.26	2.85	-	0.30	>0.05	-
		Non-TPPS	47	10.13	2.90	-			
Form 1 Unit 3 (Thermal Expansion)	Total Population	-	1146	11.25	3.45	0.64	-	-	-
	Pilot Schools	TPPS	480	11.71	3.56	-	0.4	>0.05	-
		Non-TPPS	230	11.60	3.41	-			
	Control Schools	TPPS	74	11.27	2.72	-	3.23	<0.01	0.22
		Non-TPPS	126	9.98	2.89	-			

membership of 0.22, indicating that about 4 per cent of the variance is associated with membership of the ex-TPPS and non-TPPS groups (see Table 1). The reliability for this test was much higher than for the other two tests.

Results on items related to TPPS activities. For each item on the tests which was directly related to a specific lesson in Phase 3 TPPS, comparisons were made between ex-TPPS and non-TPPS students on the basis of choice distribution and percentage correct, using Chi-square.

In the Unit 1 test, five items (1, 2, 3, 4 and 14) related to Phase 3B lessons as follows:

Lesson 1: Let's find out what the marks on a syringe mean;
Lesson 2: Let's make a solid cubic centimetre;
Lesson 3: Let's mark volumes on a glass jar;
Lesson 4: Let's record the volume of air in our lungs.

A further six items (6, 7, 8, 9, 11 and 12) were closely related to Phase 3A lessons, namely:

Lesson 4: Let's make another weighing machine;
Lesson 7: Let's make a metric tape measure;
Lesson 8: Let's study a live frog or toad;
Lesson 9: Let's study a live grasshopper;
Lesson 10: Let's study human beings.

Lessons 8, 9 and 10 all use the application of measurement skills in making observations on the organisms involved.

Of the five comparisons made on items related to Phase 3A and the six on Phase 3B related items in pilot schools, only one comparison between ex-TPPS and non-TPPS pupils (item 8, which shows a diagram of a small object held on to a rule scaled in millimetres by caliper-like blocks and asks for the length of the object illustrated) showed significance at the 0.05 level or better, with the ex-TPPS group having a higher percentage correct (38%) than the non-TPPS group (22%) (See Tables 2(a) and 2(b)).

Of the 11 similar comparisons made in the control schools, two of the items (2, 7 and 14) showed significant superiority at the 0.05 level or better for the ex-TPPS groups while one showed a reverse trend. These results are summarized in Tables 2(a) and 2(b).

In all of these cases, the Contingency Coefficients show that the strength of the difference is small (Siegel, 1956, p. 147).

In the Unit 2 test, only 3 items (18, 19 and 21) were closely related with TPPS lessons, these being the Phase 3B lessons:

Lesson 5: Let's see what substances dissolve in water;
Lesson 6: Let's compare how soluble things are;
Lesson 7: Let's get pure water from salty water;
Lesson 8: Let's mix copper sulphate and sand and separate them again.

However some of these items showed significant differences (see Table 2(c)).

The Unit 3 test contained no items directly related with TPPS Phase 3 lessons.

Conclusions

Caution has to be exercised in drawing any wide generalizations from these results. The

TABLE 2(a)

Analysis of items related to TPPS Phase 3A Lessons 4, 7, 8, 9 ,10
(Test Form 1, Unit 1)

Item	Sample	Group	Choice Distribution					% Correct	χ^2	p<	C
			A	B	C	D	Omit				
6	Pilot	TPPS	39	70	485	112	4	68	3.96	n.s.	-
		Non-TPPS	3	15	65	18	0	66			
	Control	TPPS	13	16	178	61	6	65	5.05	n.s.	-
		Non-TPPS	1	6	33	16	3	56			
7	Pilot	TPPS	72	341	243	53	1	34			
		Non-TPPS	8	50	37	4	0	37	2.45	n.s.	-
	Control	TPPS	27	125	98	19	5		8.30	0.08	-
		Non-TPPS	3	19	30	4	3				
	Control	TPPS	-	-	-	-	-	36	4.64	0.03*	0.11
	% Correct	Non-TPPS	-	-	-	-	-	51			
8	Pilot	TPPS	104	79	110	408	9		11.97	0.02*	0.12
		Non-TPPS	18	22	14	44	1				
	Pilot	TPPS						57	5.8		
	% Correct	Non-TPPS						44		0.02*	0.05
	Control	TPPS	50	35	64	120	5	44	6.22	n.s.	-
		Non-TPPS	7	8	20	21	3	36			
9	Pilot	TPPS	329	141	137	100	3	46	3.21	n.s.	
		Non-TPPS	40	26	17	15	1	40			
	Control	TPPS	117	51	46	56	4	43	7.38	n.s.	-
		Non-TPPS	21	11	15	9	3	36			
11	Pilot	TPPS	120	421	83	79	7	17	2.64	n.s.	-
		Non-TPPS	22	56	11	10	0	22			
	Control	TPPS	51	149	27	42	5	19	6.34	n.s.	-
		Non-TPPS	10	23	9	14	3	17			
12	Pilot	TPPS	33	82	96	496	3	70	8.43	n.s.	
		Non-TPPS	9	8	11	69	2	70			
	Control	TPPS	20	31	48	170	5	62	6.34	n.s.	-
		Non-TPPS	7	9	7	33	3	56			

TABLE 2(b)

Analysis of items related to TPPS Phase 3B Lessons 1, 2, 3, 4
(Test Form 1, Unit 1)

Item	Sample	Group	Choice Distribution					% Correct	χ^2	p<	C
			A	B	C	D	Omit				
1	Pilot	TPPS	172	274	205	51	8	24	2.82	n.s.	-
		Non-TPPS	25	31	36	6	1	25			
	Control	TPPS	56	126	73	12	5	20	3.22	n.s.	-
		Non-TPPS	15	23	14	4	1	25			
2	Pilot	TPPS	154	450	76	25	5	63	5.03	n.s.	-
		Non-TPPS	22	57	11	8	1	58			
	Control	TPPS	117	103	22	27	5		11.22	0.02*	0.18
		Non-TPPS	25	13	11	7	3	-			
	Control % Correct	TPPS						38	5.2	0.02*	0.12
		Non-TPPS						22			
3	Pilot	TPPS	392	215	49	47	7	55	0.71	n.s.	-
		Non-TPPS	54	33	6	5	1	55			
	Control	TPPS	107	135	11	15	6	39	2.74	n.s.	
		Non-TPPS	22	28	4	2	3	37			
4	Pilot	TPPS	255	44	291	114	6	41	3.66	n.s.	-
		Non-TPPS	37	4	47	10	1	47			
	Control	TPPS	94	29	93	53	5	34	8.1	n.s.	-
		Non-TPPS	13	9	19	14	4	32			
14	Pilot	TPPS	173	148	256	129	4	36	2.12	n.s.	-
		Non-TPPS	23	17	35	23	1	35			
	Control	TPPS	67	53	93	57	4		12.26	0.02*	0.19
		Non-TPPS	17	11	10	18	3				
	Control % Correct	TPPS						34	6.49	0.01*	0.13
		Non-TPPS						17			

TABLE 2(c)

Analysis of items related to TPPS Phase 3B Lessons 5, 6, 7, 8 (Test Form 1, Unit 2)

Item	Sample	Group	Choice Distribution					% Correct	χ^2	p	C
			A	B	C	D	Omit				
18	Pilot	TPPS	270	93	105	202	7	40	2.00	n.s.	-
		Non-TPPS	16	9	8	13	1	34			
19	Pilot	TPPS	51	539	50	34	3	80	4.3	n.s.	-
		Non-TPPS	6	39	1	1	0	83			
21	Pilot	TPPS	172	49	378	68	10	56	3.52	n.s.	-
		Non-TPPS	7	4	30	6	0	6			

* Significant at 0.05 level or better

study did not use a rigorously drawn up research design, but developed as an adjunct to the evaluation program in the piloting of new secondary science materials. However the sample can be considered as reasonably representative of Papua New Guinea form I high school students as a whole, with urban and rural, mission and government, and coastal and highland components. It was not possible to use definitive records to identify students in the sample who had been subjected to TPPS in their primary schooling, although there is no reason to suspect any significant degree of error in the categorization obtained from the method used.

There is no evidence from the results obtained for any strong impact of TPPS on achievement at Form 1 secondary level in the areas covered by the unit tests.

On the basis of whole test results, there is a slender indication that ex-TPPS students may have performed a little better in schools using the old interim syllabus, with the difference between the ex-TPPS and non-TPPS group of 0.79 in favour of the ex-TPPS group being significant at the 0.06 level in Unit 1 and a difference of 1.3 in a similar direction for the same group on Unit 3 being significant at the 0.01 level, but point biserial correlations indicate that the relationship between scores and group membership is very small.

There is a little support for this conclusion at least in the area of measurement, where two of the six items related to TPPS work on the Unit 1 test on measurement in the control schools showed a small significant difference in favour of the ex-TPPS group although one item in the test showed a reverse trend.

If this conclusion is correct, then the new secondary science syllabus Form 1 level may be working slightly better than the old interim syllabus on some of the basic concepts, despite the fact that the evaluation of pilot materials showed success levels much below those desired (Department of Education, Papua New Guinea, Secondary Science Evaluation Report, 1973). However, if the hopes of the course developers were being realized, much stronger evidence should have been produced by this study, despite its methodology deficiencies.

It is unfortunate that there can be no opportunity to explore this issue further, as TPPS is now officially to be taught in all primary schools, and there will be no longer any substantial groups of primary school students not being taught TPPS, although Wilson's evaluation indicates

a reluctance on the part of some teachers to carry out this policy (Wilson, 1974, p. 19).

Wilson's evaluation suggests that the Phase 3 TPPS is experiencing problems. The results of this small study are in line with this finding and indicate that if TPPS has had any effect at secondary level at all it is a meagre one.

There is a need for continued rigorous research into curriculum development and related achievement in Papua New Guinea. There have been many well intentioned, well thought out curriculum projects which in the long run have not appeared to be as successful as hoped, but there has been very little well designed specific evaluative research to confirm or deny such misgivings or to diagnose remedial needs.

The Three Phase Primary Science Course is a course which has been developed from well-intentioned, well thought out premises with the aid of strong resources and in theory represents an important step forward in the Papua New Guinea educational scene. Continued research into its progress is needed to ensure that the good intentions and the manpower and financial resources put into it have not been wasted.

References

COHEN, J. "Some statistical issues in psychological research", in Wolman, B.B. (ed.), *Handbook of Clinical Psychology*. McGraw-Hill, New York, 1965.

DEPARTMENT OF EDUCATION, PAPUA NEW GUINEA. *Pilot Secondary Science Syllabus.* Konedobu, Papua New Guinea, 1972.

DEPARTMENT OF EDUCATION, PAPUA NEW GUINEA. *Secondary Science Pilot Project, Evaluation report.* Konedobu, Papua New Guinea, 1973.

SIEGEL, S. *Non parametric Statistics for the Behavioral Sciences.* McGraw-Hill, New York, 1956.

UNESCO. *Planning for science teaching improvement in Asian schools: report of a regional workshop.* UNESCO Regional Office for Education in Asia, Bangkok, 1969,

WILLIAMS, A.B. *Papua and New Guinea. Science education: curriculum development and in-service teacher training.* UNESCO, Serial No: 2549/RMO. EDS, Paris, 1971.

WILSON, M. *Three Phase Primary Science Phase I and II Evaluation.* Research Report No. 14, Teaching Methods and Materials Centre, University of Papua New Guinea, 1972(a).

WILSON, M. *Three Phase Primary Science Phase III Evaluation. Preliminary Report.* Research Report No. 18, Teaching Methods and Materials Centre, University of Papua New Guinea, 1972(b).

WILSON, M. *Three Phase Primary Science Phase Three Evaluation: Interim Report.* Research Report No. 24, Teaching Methods and Materials Centre, University of Papua New Guinea, 1974.

SOME ISSUES IN THE SIMULATION OF ANOVA BY MULTIPLE REGRESSION TECHNIQUES

John T. McArthur, Barry J. Fraser & Leo H.T. West

Introduction

A major contribution to research in psychology and education was made by Cohen (1968) in a clear and convincing demonstration of the equivalence of analysis of variance (AV) and multiple regression (MR) techniques.

Cohen's appealing thesis consisted of two simple points. Firstly, any AV solution can be simulated using MR. Secondly, and more importantly, MR offers a number of practical advantages not realisable within the AV framework.

Fundamentally, MR differs from AV in terms of the added flexibility it offers the researcher. For example, within the AV framework, any continuous independent variable like IQ must be reduced to a number of "pseudo-qualitative" categories (Hurst, 1970, p. 541). But Cohen (1968, p. 441) has claimed that this "wilful degradation of available measurement information has a direct consequence in the loss of statistical power", while Kerlinger and Pedhazur (1973, Ch. 10) have provided numerical examples illustrating this loss of power. By permitting the use of continuous predictor variables, MR overcomes this important limitation present with AV.

MR not only allows the inclusion of the same interactions which are found with AV but, also, permits the option of excluding some interactions when desired. This point has been graphically illustrated in Cohen's following remarks:

> The textbook paradigms for factorial design AV lead data analysts to dutifully harvest all possible interactions of all possible orders up to the highest, whether or not they are meaningful or interpretable or, if interpretable, communicable. There emanate from psychology departments many silent prayers to the spirit of R.A. Fisher that high order interactions will not prove significant! (Cohen 1968, p. 438)

AV computer programs often lack the capacity to analyze data for a wide range of experimental designs containing large numbers of predictor variables, factors with multiple levels or varying numbers of covariates. With MR, however, McArthur and West (1974) have pointed out that the one general purpose MR program would be capable of handling quite large numbers and varieties of variables.

MR also offers flexibility in other ways. It readily permits the use of "preliminary combined tests . . . prior to doing individual tests" (Appelbaum & Cramer, 1974) and the blocks of predictor variables used by Fraser (1975). MR also allows the analysis of curvilinear terms (Walberg, 1971). As well, MR analyses automatically yield estimates of the amount of variance explained by predictors (Kerlinger, 1973, p. 618). [It should be noted, however, that analogous indices are available, although not commonly used, with AV (Schutz, 1966; Vaughan & Corballis, 1969; Halderson & Glasnapp, 1972)]. Furthermore, MR offers a variety of solutions, to be discussed in more detail later, for the analysis of data from unequal cells designs.

Despite the persuasiveness of Cohen's publication, and the flurry of literature which has followed it, the practice of science education research seems to have been hardly affected. The use of AV remains entrenched while the use of MR is relatively rare. Even those few articles employing MR techniques often display serious flaws. Grobe's article (1973), for example, exhibited several misunderstandings of MR which have been discussed by McArthur (in press).

While Cohen was successful in achieving his major purpose of presenting multiple regression as a general data analytic tool, two sins of omission can be identified. The first was that, while arguments for the equivalence of AV and MR were convincing, explication of the actual methodology of achieving the simulation was too brief to enable unambiguous interpretation and use. Similarly, McArthur and West (1974) have demonstrated a method for the simulation of a two-way AV, but this method is not readily generalizable to more complex designs. Cohen also failed to discuss adequately the problems associated with nonorthogonal designs. In fact, Cohen's basic thesis – that MR could simulate ordinary AV with balanced data and that, as well, offered advantages such as providing solutions with unbalanced data – has led many people to the erroneous belief that the interpretation placed on balanced AV can be transferred to the unbalanced design. But this is not the case. With unbalanced designs, there exist a variety of different solutions and interpretations.

It is the purpose of the remainder of this paper to attempt to overcome the two omissions in Cohen's article by providing some explication of methods for simulating AV, and by raising a number of important issues which surround the use of MR to provide solutions for unbalanced data.

Simulation of equal cells AV

In order to demonstrate methods of simulating AV using MR, attention will first be given to a special AV design, namely, the equal cells factorial design. This design is considered as a special case because, for equal cells AV, all main effects and interactions are *mutually orthogonal*. Therefore in attempting to simulate equal cells AV using MR, it is essential to ensure that orthogonality exists between all the terms of the regression equation. In practice, orthogonality between terms in a regression model can be achieved by the use of an appropriate method of *coding* of categorical predictors together with the choice of an appropriate *algorithm* for the computation of the F statistic.

Table 1 provides data for a three-way equal cells factorial design taken from Veldman (1967, p. 234), which will be used to illustrate the method of MR simulation. The table also shows two different ways of coding each categorical main effect and interaction for use in a regression solution. The first scheme, *dummy coding*, employs 1's and 0's to designate each of the levels of the dichotomous variables while the second scheme, *effect coding*, uses 1's and -1's. In both coding schemes, the code for an interaction is simply a multiplicative function of the codes for the corresponding main effects. Further discussion of coding may be found in Suits (1957) and Kerlinger and Pedhazur (1973).

As well as alternatives in coding, MR analysis allows the use of a variety of different algorithms for computing F. Of these different algorithms, the three identified by Overall and Spiegel (1969) have received particular attention in the literature. Using a two-way AV as an example, these three algorithms have been illustrated algebraically by Overall and Spiegel (1969, p. 316) and Overall and Klett (1972, p. 448) and diagramatically by Kaufman and Sweet (1974, p. 366). In the literature, however, little guidance has been given in generalizing from the two-way case to more complex designs.

For this reason, Table 2 has been compiled to illustrate the use of each of the three Overall and Spiegel methods in computing the F statistic for all seven terms present in a three-way AV.

In Table 2, R simply represents the multiple correlation between the criterion and a set of predictors while R^2 represents the amount of variance explained. For example, R_{123} would

TABLE 1

Two coding methods for a 2 x 2 x 2 AV

Cell			Y	Dummy coding							Effect Coding						
				X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_1	X_2	X_3	X_4	X_5	X_6	X_7
				A	B	C	AB	AC	BC	ABC	A	B	C	AB	AC	BC	ABC
A_1	B_1	C_1	25	1	1	1	1	1	1	1	1	1	1	1	1	1	1
			26	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		C_2	20	1	1	0	1	0	0	0	1	1	−1	1	−1	−1	−1
			30	1	1	0	1	0	0	0	1	1	−1	1	−1	−1	−1
	B_2	C_1	23	1	0	1	0	1	0	0	1	−1	1	−1	1	−1	−1
			28	1	0	1	0	1	0	0	1	−1	1	−1	1	−1	−1
		C_2	28	1	0	0	0	0	0	0	1	−1	−1	−1	−1	1	1
			25	1	0	0	0	0	0	0	1	−1	−1	−1	−1	1	1
A_2	B_1	C_1	26	0	1	1	0	0	1	0	−1	1	1	−1	−1	1	−1
			28	0	1	1	0	0	1	0	−1	1	1	−1	−1	1	−1
		C_2	24	0	1	0	0	0	0	0	−1	1	−1	−1	1	−1	1
			26	0	1	0	0	0	0	0	−1	1	−1	−1	1	−1	1
	B_2	C_1	30	0	0	1	0	0	0	0	−1	−1	1	1	−1	−1	1
			29	0	0	1	0	0	0	0	−1	−1	1	1	−1	−1	1
		C_2	29	0	0	0	0	0	0	0	−1	−1	−1	1	1	1	−1
			29	0	0	0	0	0	0	0	−1	−1	−1	1	1	1	−1

TABLE 2

Overall and Spiegel's three algorithms for computing *F*

Effect	Method 1	Method 2	Method 3
1. A	$\frac{(R^2_{1234567} - R^2_{234567})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{123} - R^2_{23})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{1})/dfn}{(1 - R^2_{1234567})/dfd}$
2. B	$\frac{(R^2_{1234567} - R^2_{134567})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{123} - R^2_{13})dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{12} - R^2_{1})/dfn}{(1 - R^2_{1234567})/dfd}$
3. C	$\frac{(R^2_{1234567} - R^2_{124567})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{123} - R^2_{12})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{123} - R^2_{12})/dfn}{(1 - R^2_{1234567})/dfd}$
4. AB	$\frac{(R^2_{1234567} - R^2_{123567})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{1234} - R^2_{123})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{1234} - R^2_{123})/dfn}{(1 - R^2_{1234567})/dfd}$
5. AC	$\frac{(R^2_{1234567} - R^2_{123467})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{1235} - R^2_{123})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{12345} - R^2_{1234})/dfn}{(1 - R^2_{1234567})/dfd}$
6. BC	$\frac{(R^2_{1234567} - R^2_{123457})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{1236} - R^2_{123})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{123456} - R^2_{12345})/dfn}{(1 - R^2_{1234567})/dfd}$
7.ABC	$\frac{(R^2_{1234567} - R^2_{123456})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{1234567} - R^2_{123456})/dfn}{(1 - R^2_{1234567})/dfd}$	$\frac{(R^2_{1234567} - R^2_{123456})/dfn}{(1 - R^2_{1234567})/dfd}$

R multiple correlation; dfn degrees of freedom in numerator; dfd degrees of freedom in denominator

represent the multiple correlation of the criterion with the A, B and C effects, while $R^2_{1234567}$ would indicate the amount of variance explained by the full model containing all seven predictor terms present in a three-way AV. The number of degrees of freedom in the numerator, *dfn*, is equal to the difference between the number of predictors in the larger model in the numerator and the number of predictors in the smaller model. The number of degrees of freedom in the denominator, *dfd*, is equal to the number of subjects, less one, minus the number of predictors contained in the full model which appears in the denominator.

In Method 1, each effect, whether it be main effect or interaction, is adjusted for relationship to all other effects in the model. Method 2, which has been called the Experimental Design Approach by Kerlinger and Pedhazur (1973) and the Method of Fitting Constants by Williams (1972), adjusts each effect for all other effects at a lower level, ignoring higher order effects. For Method 3, also known as the A Priori Ordering Approach (Kerlinger & Pedhazur, 1973) or the Hierarchical Model (Williams, 1972), an ordering is assumed for all terms and each term is adjusted only for those preceding it in the ordering. For all three methods, the same error term is used, as suggested by Kaufman and Sweet (1974) and Kempthorne (1952).

Data for the equal cells design in Table 1 has been used in an attempt to simulate AV results using both Method 1 and Method 2. Table 3 shows the AV results obtained by Veldman (1967, p. 226) for this problem, and four different MR solutions – Method 1 with dummy coding, Method 1 with effect coding, Method 2 with dummy coding and Method 2 with effect coding. Of the four MR solutions, only three gave *F* values consistent with AV results. The MR solution which failed to simulate AV results was Method 1 used with dummy coding. The failure of Method 1 with dummy coding is also indicated by the fact that the solution was not of full rank; that is, the sum of the R^2 values for all terms was less than the value of R^2 for the full model.

In an attempt to understand the failure of Method 1 with dummy coding, reference can be made to the intercorrelation matrices in Table 4. This table shows that, while all correlations between predictors are zero with effect coding, some non-zero correlations exist when dummy coding is employed. That is, although main effects are mutually orthogonal using dummy coding, interaction terms are nonorthogonal with both the main effects and the interaction containing component terms.

Because equal cells AV is a special case in which all terms are mutually orthogonal, MR simulation can be successful only if all terms in the regression model can be rendered mutually orthogonal also. As Table 4 shows, effect coding can be used to achieve this orthogonality. Furthermore, provided effect coding is used, simulation of AV would be successful irrespective of which algorithm is employed. On the other hand, successful simulation using dummy coding would only be possible with algorithms which exploited the use of regression models containing uncorrelated terms, as in Method 2.

The failure of Method 1 with dummy coding to simulate AV results is of great importance. Yet this point has been often misunderstood in the literature. Veldman (1967), for example, has recognized the potential use of MR for AV simulation and has provided the computer program REGRAN for performing MR analyses. However, Veldman's program was written only for Method 1 and his discussion of coding was restricted entirely to dummy coding. It is tempting, therefore, for the researcher to employ Method 1 and dummy coding to achieve the erroneous results seen earlier.

Proportional cell case

In the equal cells case, discussed in the previous section, it was seen that the use of effect

TABLE 3

Five solutions to a three-way equal cells AV

Predictor	AV Solution	MR Method 1				MR Method 2			
		Dummy Coding		Effect Coding		Dummy Coding		Effect Coding	
	F	R^2	F	R^2	F	R^2	F	R^2	F
A	1.778	0.0541	0.695	0.1382	1.778	0.1382	1.778	0.1382	1.778
B	1.778	0.1382	1.777	0.1382	1.778	0.1382	1.778	0.1382	1.778
C	0.111	0.0021	0.027	0.0086	0.111	0.0086	0.111	0.0086	0.111
AB	0.694	0.0271	0.348	0.0540	0.694	0.0540	0.694	0.0540	0.694
AC	0.250	0.0097	0.124	0.0194	0.250	0.0194	0.250	0.0194	0.250
BC	0.250	0.0098	0.126	0.0194	0.250	0.0194	0.250	0.0194	0.250
ABC	0.000	0.0000	0.000	0.0001	0.001	0.0000	0.000	0.0001	0.001
ΣR^2		0.2410		0.3779		0.3778		0.3779	
Full Model R^2		0.3779		0.3780		0.3779		0.3780	

TABLE 4

Intercorrelaton matrices

	Dummy Coding							Effect coding					
	B	C	AB	AC	BC	ABC		B	C	AB	AC	BC	ABC
A	.00	.00	.58	.58	.00	.38	A	.00	.00	.00	.00	.00	.00
B		.00	.58	.00	.58	.38	B		.00	.00	.00	.00	.00
C			.00	.58	.58	.38	C			.00	.00	.00	.00
AB				.33	.33	.65	AB				.00	.00	.00
AC					.33	.65	AC					.00	.00
BC						.65	BC						.00

coding in equal cells designs yielded zero correlation between all terms of an AV design and that Method 1 and Method 2 both simulated AV solutions. It can be readily seen that Method 3 would also yield identical results in this case provided that the appropriate *a priori* ordering was specified; that is, main effects should precede interactions and lower order interactions should precede higher order interactions.

In the proportional cell case, however, some additional problems arise. Even when effect

coding is used, only the main effects are orthogonal while correlations would exist amongst interactions and between main effects and interactions (Gorsuch, 1973, p. 90). Because of this correlation, Overall and Spiegel (1969) found that Method 1 results were inconsistent with the proportional cell AV solutions of Lindquist (1953, p. 228), whereas Method 2 and Method 3 results were consistent. On the basis of these analyses, Overall and Spiegel (1969, p. 318) concluded that, for proportional data, "Method 2 is the proper generalization of traditional experimental analysis of variance and Method 1 is something else". This same conclusion, based upon an attempted generalization of the equal cells method of solution to the proportional case, has been repeated without critical review in Overall and Klett (1972) and Williams (1972).

Gocka (1973), however, has demonstrated that Method 1 does in fact give results consistent with AV solutions provided that appropriate coding is used. He showed that ordinary effect coding (using 1, −1 or 1, 0, −1) was unsuitable in the proportional cell case and advocated a different coding scheme which took account of variation in cell size. Using this new coding scheme, shown in Gocka (1973, p. 27), it was demonstrated that Method 1 did simulate AV results in the proportional cell case.

Disproportional cell designs

So far, discussion has been restricted to equal cells and proportional cells designs. In educational and psychological research, however, naturalistic data are disproportional more often than not according to Kaufman and Sweet (1974, p. 364), Fennessey (1969, p. 27) and Kerlinger and Pedhazur (1973, p. 77). But disproportional cell designs are very different from either the equal or proportional case. In disproportional cells designs, non-zero correlation exists between all terms so that there is overlap between the variances attributable to each effect. And, furthermore, this overlap cannot meaningfully be removed solely by the choice of coding as in the equal cells and proportional cells designs. For the researcher employing designs, the major problem is to decide how the overlapping variance is to be partitioned among the different predictor variables.

Our problem with nonorthogonal AV designs, therefore, is not merely that of finding an appropriate method for simulation by MR. As well, the AV solution itself "remains inherently ambiguous" (Goldberger, 1964, p. 201) because "the number of different ways of organizing the overlapping variances are infinite" (Gorsuch, 1973, p. 100). The problem of partitioning overlapping variance, then, is common to both the AV and MR approaches. Once this question is resolved, MR simulation can be accomplished using appropriate coding and algorithms as discussed earlier.

Overall and Spiegel (1969, p. 311) have issued a clear warning that "researchers should be cautioned against placing analysis of variance interpretations on tests of hypotheses that are not like those provided by the usual balanced experimental design". It is confusing to find statements such as that by Jennings (1967, p. 105) who claims that "the logic is unchanged when there are an unequal number of cases per cell although the form the vectors take may be surprising".

Not only is the interpretation of unbalanced data different from that of balanced data, but also there exists a variety of different solutions to any unbalanced design. In addition to the three methods of Overall and Spiegel (1969), described earlier, Rawlings (1972 & 1973) has proposed a fourth method and Williams (1972) has identified another six solutions. As there are at least 10 different solutions to any unbalanced design, the researcher must choose the most appropriate method according to "the theoretical perspective of the study" (Gorsuch, 1973, p. 89) and "research-strategic issues that go to the heart of the substantive nature of the research" (Cohen, 1968, p. 442).

Conclusion

For the equal and proportional cell cases, simple procedures (that are applicable to complex as well as simple designs) can be used to simulate AV using MR. Overall and Spiegel's Method II algorithm can always be used. Their Method I algorithm can be used for the equal cells case provided that effect coding is used (i.e. for a two factor effect the coding 1, −1 is used, for an n factor effect the coding 1, 0, 0, ... , −1 is used). For the proportional cell case, Gocka's alternative coding needs to be substituted for effect coding.

MR (or "least squares solution") does not "solve" the disproportional cell case. There is no one-to-one correspondence between equal and proportional cell and disproportional cell AV. An AV-type solution depends on re-definition of the main effect variables to orthogonalize them. Much of the confusion about which MR solution to use would be removed if researchers decided on logical grounds how they wished to re-define the main effect variables. The choice of the appropriate MR procedure is then a straightforward one.

References

APPELBAUM, M.I. & CRAMER, E.M. "Some problems in nonorthogonal analysis of variance." *Psychological Bulletin*, **81**: 335-343, 1974.

COHEN, J. "Multiple regression as a general data-analytic system." *Psychological Bulletin*, **70**: 426-443, 1968.

FENNESSEY, J. "The general linear model: A new perspective on some familiar topics." *American Journal of Sociology*, **74**: 1-27, 1968.

FRASER, B.J. "The impact of ASEP on pupil learning and classroom climate." *Research in Science Education*, 1975, **5**, (this issue).

GOCKA, E.F. "Regression analysis of proportional cell data." *Psychological Bulletin*, **80**: 25-27, 1973.

GOLDBERGER, A.S. *Economic Theory*. John Wiley, New York, 1964.

GORSUCH, R.L. "Data analysis of correlated independent variables." *Multivariate Behavioral Research*, **8**: 89-107, 1973.

GROBÉ, C.H. "A regression approach to evaluating instructional programs in science." *Journal of Research in Science Teaching*, **10**: 55-62, 1973.

HALDERSON, J.S. & GLASNAPP, D.R. "Generalized rules for calculating the magnitude of an effect in factorial and repeated measures ANOVA designs." *American Educational Research Journal*, **9**: 301-310, 1972.

HURST, R.L. "Qualitative variables in regression analysis." *American Educational Research Journal*, **7**: 541-552, 1970.

JENNINGS, E. "Fixed effects analysis of variance by regression analysis." *Multivariate Behavioral Research*, **1**: 95-107, 1967.

KAUFMAN, D. & SWEET, R. "Contrast coding in least squares regression analysis." *American Educational Research Journal*, **11**: 359-377, 1974.

KEMPTHORNE, O. *The Design and Analysis of Experiments*. John Wiley, New York, 1952.

KERLINGER, F.N. & PEDHAZUR, E.J. *Multiple Regression in Behavioral Research*. Holt, Rinehart and Winston, New York, 1973.

KERLINGER, F.N. *Foundations of Behavioral Research*. Second edition, Holt, Rinehart and Winston, New York, 1973.

LINDQUIST, E.F. *Design and Analysis of Experiments in Psychology and Education.* Houghton Mifflin, Boston, 1953.

MCARTHUR, J.T. & WEST, L.H.T. "Multiple regression as a general data analysis technique." *Research in Science Education,* **4**: 185-192, 1974.

MCARTHUR, J.T. "Some comments on 'A regression approach to evaluating instructional programs in science'." *Journal of Research in Science Teaching* (in press).

OVERALL, J.E. & SPIEGEL, D.K. "Concerning least squares analysis of experimental data." *Psychological Bulletin,* **72**: 311-322, 1969.

OVERALL, J.E. & KLETT, C.J. *Applied Multivariate Analysis.* McGraw-Hill, New York, 1972.

RAWLINGS, R.R., Jr. "Note on nonorthogonal analysis of variance." *Psychological Bulletin,* **77**: 373-374, 1972.

RAWLINGS, R.R., Jr. "Comments on the Overall and Spiegel paper." *Psychological Bulletin,* **79**: 168-169, 1973.

SCHUTZ, R.E. "The control of 'error' in educational experimentation." *The School Review,* **74**: 150-158, 1966.

SUITS, D.B. "Use of dummy variables in regression equations." *Journal of the American Statistical Association,* **52**: 548-551, 1957.

VAUGHAN, G.M. & CORBALLIS, M.C. "Beyond tests of significance: Estimating strength of effects in selected ANOVA designs." *Psychological Bulletin,* **72**: 204-213, 1969.

VELDMAN, D.J. *Fortran Programming for the Behavioral Sciences.* Holt, Rinehart and Winston, New York, 1967.

WALBERG, H.J. "Generalized regression models in educational research." *American Educational Research Journal,* **8**: 71-91, 1971.

WILLIAMS, J.D. "Two way fixed effects analysis of variance with disproportionate cell frequencies." *Multivariate Behavioral Research,* **1**, 67-83, 1972.

THE CONTRIBUTORS

A.H.A. BENSINK	Department of Entomology, University of Queensland.
EFFIE D. BEST	Research and Planning Branch, Education Department, South Australia.
P.R.B. BLOOD	Department of Entomology, University of Queensland.
BRIAN W. CARSS	Education Department, University of Queensland.
BRIAN G. COLLINS	Department of Biology, Western Australian Institute of Technology.
PETER FENSHAM	Faculty of Education, Monash University.
JOHN W. FISHER	Murray Park College of Advanced Education, South Australia.
BARRY J. FRASER	Faculty of Education, Monash University.
P.L. GARDNER	Faculty of Education, Monash University.
W.C. HALL	Advisory Centre for University Education, University of Adelaide.
E. JUNGWIRTH	Hebrew University of Jerusalem, Israel.
JANET V. KEIGHTLEY	Research and Planning Branch, Education Department, South Australia.
ROBERT L. KIMBER	Cleve Area School, South Australia.
J. LAWRENCE	Queen Elizabeth Hospital, Woodville, South Australia.
LINDSAY D. MACKAY	Faculty of Education, Monash University.
JOHN T. McARTHUR	Faculty of Education, Monash University.
M.N. MADDOCK	Department of Education, University of Newcastle, New South Wales.
B.C. NEWMAN	Faculty of Education, University of New South Wales.
KAYE NICKLESS	Faculty of Education, Monash University.
JEFF R. NORTHFIELD	Faculty of Education, Monash University.
JOHN M. OWEN	Australian Council for Educational Research, Hawthorn, Victoria.
COLIN N. POWER	Department of Education, University of Queensland
R.J. STANTON	District Education Office, Mt. Hagen, Papua New Guinea.
R.P. TISHER	Faculty of Education, Monash University.
R. TREMBATH	State College of Victoria, Frankston.

LEO H. T. WEST Higher Education Advisory & Research Unit, Monash University.

R.T. WHITE Faculty of Education, Monash University.